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 8—1974

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

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

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

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 64-8029

ISBN 0-12-034808-X

PRINTED IN THE UNITED STATES OF AMERICA

LIST OF CONTRIBUTORS

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

G. BERTHIER, Laboratoire de Chimie Quantique, Institut de Biologie Physico-Chimique, Paris, France (183)

GIUSEPPE DEL RE, Cattedra di Chimica Teorica, Istituto di Chimica dell' Università, Napoli, Italy (95)

CHRISTIAN KLIXBÜLL JØRGENSEN, Département de Chimie Physique, Université de Genève, Geneva, Switzerland (137)

DAVID A. MICHA, Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida (231)

WILLIAM I. SALMON,* Ames Laboratory of the United States Atomic Energy Commission and Department of Chemistry, Iowa State University, Ames, Iowa (37)

J. SERRE, Laboratoire de Chimie ENSJF, 92120 Montrouge, France (1)

---

* *Present address*: Department of Physics, University of Utah, Salt Lake City, Utah.

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 ninth 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 covered in this volume range from the treatment of symmetry properties over molecular orbital studies involving the Hartree–Fock method and photoelectron spectra to the theory of molecular collisions. Some of the papers emphasize studies in fundamental quantum theory, 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

Symmetry Groups of Nonrigid Molecules

J. SERRE
Laboratoire de Chimie ENSJF
92120 Montrouge, France

I. Introduction

In recent years, molecules with nonrigid molecular skeletons have attracted wide interest. The properties of such molecules have been extensively studied due to the development of powerful experimental techniques. A molecule is said to be nonrigid (Longuet-Higgins, 1963) if it is in an electronic state that has several potential energy minima separated by surmountable energy barriers. If these barriers are very small compared to kT (i.e., around 600 cal/mole at 300°K), rapid isomeric change will occur and one part of the molecule will rotate freely with respect to the other. Barrier heights are determined mainly by kinetic methods, by infrared and Raman spectroscopy, by microwave spectroscopy, and by neutron scattering (Pethrick and Wyn-Jones, 1969), and their values vary quite widely [for example, 6 cal/mole for the nitromethane molecule (Tannenbaum *et al.*, 1956), 1177 cal/mole for the methylketene molecule (Bak *et al.*, 1966), and 2590 cal/mole for the dimethyl ether molecule (Durig and Li, 1972; Tuazon and Fateley, 1971)]. But other kinds of displacement can occur: For example, the C_5 ring of the cyclopentane molecule is permanently distorted and the molecule is thus nonrigid, undergoing low-frequency large-amplitude out-of-plane displacements; in consequence, a free pseudo-rotation occurs (Mills, 1971). The monocentric molecules offer examples of rearrangements that correspond to permutations of ligands or skeletal positions by bond angle deformations at the central atom (Muetterties, 1970) and the polycentric systems offer examples of conformational interconversions (Barton, 1970) by rotation about bond axes. Different possible mechanisms (Ugi *et al.*, 1971; Gillespie *et al.*, 1971) have been proposed for these positional exchanges: In the Berry pseudo-rotation mechanism (Berry, 1960) proposed for the trigonal-bipyramidal pentacoordinate compounds, the positional exchange occurs by pairwise exchange of apical and equatorial ligands. In the turnstile rotation mechanism (Ugi *et al.*, 1971), one apical and one equatorial ligand rotate as a pair (1,4 in Fig. 1) vs. the oppositely rotating trio of the remaining ligands (2,3,5 in Fig. 1) and the rotation resembles that of a turnstile.

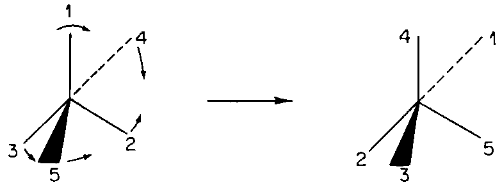


Fig. 1. Turnstile rotation mechanism.

Many experimental techniques of spectroscopy allow study of these nonrigid molecules, but the classification of the different spectroscopic states is particularly difficult; this situation results from the motions of great amplitude that take place.

In Section II of this review we present the different theories that exist for determining the symmetry group allowing classification of the molecule levels; but, before studying the nonrigid molecules more specifically, we recall how the classification of the different levels (electronic, vibrational, vibronic, rovibronic) of a rigid molecule is possible. Several experimental examples in which the different theories are applied are given in Section III. They are taken from microwave, infrared, ultraviolet, and NMR spectroscopy. Finally, in two appendices, the theorems required for building up the character tables of these particular symmetry groups are provided and the building up of such a character table is explained.

II. Theory

A. Rigid Molecules

The complete treatment of a molecule (Wilson *et al.*, 1955) must include the determination of the kinetic energy with regard to a proper coordinate system that must allow treatment of the vibrational motion independently of the rotational motion and of the translational motion of the whole molecule.

These coordinates consist of the Cartesian coordinates X, Y, Z which describe the position of the center of mass, the three Eulerian angles θ, φ, χ which describe the orientation in space of a set of rotating coordinate axes x, y, z whose origin coincides with the center of mass $0'$, and finally $(3N - 6)$ normal coordinates which give the positions of the atoms relative to each other in the rotating axis system.

It can be found in many textbooks (e.g., Wilson *et al.*, 1955) that, with such coordinates, the kinetic energy T is given by the following expression:

$$2T = \dot{\mathbf{R}}^2 \sum_\alpha m_\alpha + \sum_\alpha m_\alpha(\boldsymbol{\omega} \wedge \mathbf{r}_\alpha) \cdot (\boldsymbol{\omega} \wedge \mathbf{r}_\alpha) + \sum_\alpha m_\alpha \mathbf{v}_\alpha{}^2 + 2\boldsymbol{\omega} \cdot \sum_\alpha (m_\alpha \boldsymbol{\rho}_\alpha \wedge \mathbf{v}_\alpha),$$

where $\dot{\mathbf{R}}$ is the velocity of $0'$ in the fixed system, m_α is the mass of the particle α, $\boldsymbol{\omega}$ the angular velocity of the rotating system of axes, $\mathbf{r}_\alpha$ the instantaneous position vector of the nucleus α in the rotating system, $\mathbf{v}_\alpha$ the vector with the components x'_α, y'_α, z'_α in the rotating system, and $\boldsymbol{\rho}_\alpha$ the displacement vector of the nucleus α also in the rotating system.

The first term in the kinetic energy equation gives the translational energy, the second term, the rotational energy, and the third term, the vibrational energy. The last term represents the interaction energy between the rotation and the vibration and is called the Coriolis coupling energy. From this expression for T, it is possible to write the corresponding Hamiltonian and it will be possible to obtain a function representing the molecular motions in the form of a product of a vibrational function and a rotational function, if the Coriolis interaction is neglected.

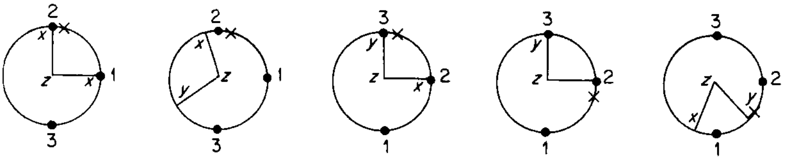
The Hamiltonian of a molecule taken on the whole may be written in the following form: $H = H_{\mathrm{e}} + H_{\mathrm{v}} + H_{\mathrm{ve}} + H_{\mathrm{r}} + H_{\mathrm{rv}}$, where H_{e} is the purely electronic part, H_{v} is the vibrational part, H_{ve} represents the vibronic interaction, H_{r} is the rotational part, and H_{rv} represents the Coriolis interaction.

The symmetry group that can be used to study a problem and to classify the corresponding wavefunctions is the one in which the elements commute with the corresponding Hamiltonian; the elements of such a group leave the considered Hamiltonian invariant.

In the constraints of the Born–Oppenheimer approximation, the electronic Hamiltonian H_e depends parametrically on the vectors a_α^0 that denote the positions of equilibrium of the atoms with respect to the molecular axes x, y, z. Let us consider the case of a triangular molecule X_3 in which the atoms are at the corners of an equilateral triangle at equilibrium (Altmann, 1971). If the axes $0x$ and $0y$ are rotated by $+2\pi/3$ about the $0z$ axis, the new values of the parameters a_α ($\bar{a}_1^0$, $\bar{a}_2^0$, $\bar{a}_3^0$) are equal to the old values a_3^0, a_1^0, and a_2^0, respectively. Then we have the following relations:

$$H(\bar{a}_1^0, \bar{a}_2^0, \bar{a}_3^0) = H(a_3^0, a_1^0, a_2^0) = H(a_1^0, a_2^0, a_3^0).$$

This symmetry operation is said to be taken in the passive convention: The nuclei and the electrons remain in fixed positions and the axes alone are moved. On the contrary, in a symmetry operation taken in the active convention (Altmann, 1962), the nuclei and the electrons are moved and the axes remain in fixed positions. In Fig. 2, the rotation C_3^+ is shown in the passive convention and the rotation C_3^- in the active convention; a permutation (132) and a Euclidean rotation R_3^- are also pictured. In a permutation (132), the nucleus 1 replaces the nucleus 3, and so forth; the electrons and the axes are fixed. In a Euclidean rotation R_3^-, the whole molecule (nuclei, electrons, axes) turns about $0z$ of $-2\pi/3$.



E $\quad$ C_3^+ passive $\quad$ (132) $\quad$ C_3^- active $\quad$ R_3^-

Fig. 2. Definitions of symmetry operations. The small circles indicate the positions of equilibrium of the atoms X; the electronic wavefunction is evaluated at the point marked with a cross.

Let us now consider the case of the molecule X_3 vibrating with the atoms at instantaneous positions r_α and define displacement vectors $\rho_\alpha = r_\alpha - a_\alpha^0$. In this part of the work, we will ignore the vibronic interactions as well as the rotational effects. The corresponding Hamiltonian can be written as

$$H = H_e(a_1^0, a_2^0, a_3^0) + H_v(\rho_1, \rho_2, \rho_3).$$

The symmetry operations acceptable for H_v must leave a_i^0 (for all i) invariant in order to leave the Hamiltonian H_e identically invariant. These symmetry operations were introduced by Wigner (1930) and they can be defined as follows: (a) the position vectors of the nuclei undergo a symmetry operation of the point group, each nucleus going with its displacement vector; (b) the nuclei are then relabeled in such a way that the original labeling is maintained.

In Fig. 3 the differences between an operation $C_3{}^+$ of the point group and a Wigner's operation $\bar{C}_3{}^+$ are shown.

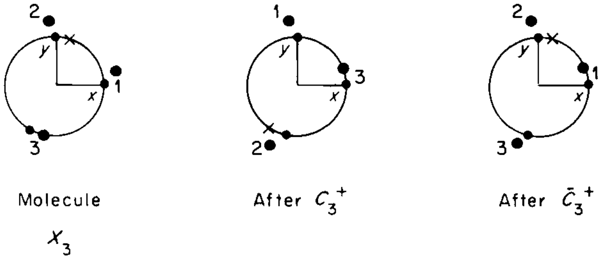


Fig. 3. Point group operation and Wigner symmetry operation. The small circles indicate the positions of equilibrium of the atoms X and the large circles their instantaneous positions.

We will now consider the molecule X_3 without neglecting the interaction between the electronic motions and the vibrational motions. If these interactions are taken into account, a Wigner's operation is not a symmetry operation of the total Hamiltonian because the relation between the value of the electronic wavefunction in one point and the displacement of a given nucleus is not preserved by this operation. Hougen (1962, 1963) introduced a new type of operation in which the vibronic interactions are kept constant; these Hougen's operations leave the Hamiltonian invariant:

$$H = H_e + H_v + H_{ve}.$$

Such an operation can be defined as follows: It is a Wigner's operation S followed by a transformation of the wavefunction such that the value of the electronic part of the wavefunction at the point Sr equals its old value at the point r. It is possible to define this operation in a different way: It is possible to associate with an operation S of the symmetry group taken in the active convention a permutation P_S that expresses the relabeling of the nuclei. For example, the permutation P_S associated with the operation $C_3{}^+$ (Fig. 3) is the circular permutation (123) where the nucleus 1 replaces the nucleus 2, and so on. The Hougen's operation S^{H} associated with the operation S is

$$S^{\mathrm{H}} = SP_{S^{-1}} = P_{S^{-1}}S.$$

It can be shown that the group of the operations S^{H} is isomorphic to the point group S.

We now consider the total Hamiltonian

$$H_{\mathrm{t}} = H_{\mathrm{e}} + H_{\mathrm{v}} + H_{\mathrm{ve}} + H_{\mathrm{r}} + H_{\mathrm{rv}}$$

where H_{r} is the rotational Hamiltonian and where H_{rv} stands for the Coriolis interaction between the rotations and the vibrations. Hougen (1962, 1963) showed that for rigid and nonlinear molecules, it is possible to classify the rovibronic wavefunctions with the help of the permutation inversion (PI) group. This group consists of all feasible permutations P (Dalton, 1966) of the space and spin coordinates of identical nuclei in the molecule (including the identity E) and all feasible permutation inversions $P^* = PE^*$ (where E^* is the operation of inversion, in the molecular center of mass, of the space coordinates of all particles, nuclei, and electrons in the molecule). This concept of permutation inversion group introduced later by Longuet-Higgins (1963) was implicitly used by Hougen (1962). In the case of rigid and nonlinear molecules the PI group and the point group are isomorphic.

For example, in the case of the water molecule, the four operations of the point group C_{2v} are E, C_{2z}, σ_{yz}, and σ_{xz} (Fig. 4); the four operations of the PI group are E, (12), E^*, and $(12)^*$ where (12) is the permutation of the space and spin coordinates of nuclei 1 and 2, E^* is the operation of inversion in the molecular center of mass of the space coordinates of all particles, nuclei, and electrons in the molecule, and $(12)^* = (12)E^*$. The operations for the water molecule are related as follows:

$$\begin{aligned} E &\equiv R_0 E, \\ (12) &\equiv R_{2z} C_2, \\ E^* &\equiv R_{2x} \sigma_{yz}, \\ (12)^* &\equiv R_{2y} \sigma_{xz}, \end{aligned}$$

where R_0 is the operation of no rotation and R_{2x}, R_{2y}, R_{2z} are the Euclidean rotations by π around the x, y, and z axes.

Let us consider the point group operation σ_{xz}, an operation of the type Hougen defined that keeps the vibronic Hamiltonian invariant (Fig. 4). The point group operation changes the vibronic angular momentum.

In this case, for example, $(\boldsymbol{\rho} \wedge \mathbf{v})_y$ is left invariant and $(\boldsymbol{\rho} \wedge \mathbf{v})_x$ and $(\boldsymbol{\rho} \wedge \mathbf{v})_z$ change sign, but the point group operation does not affect the overall rotational angular momentum and the Coriolis coupling energy is changed by the point group operation. Thus the rovibronic Hamiltonian is not invariant to the operations of the point group. If σ_{xz} is followed by the Euclidean rotation R_{2y}, the vibronic variables are unaffected but the overall rotational angular momentum is changed and the Coriolis coupling energy is invariant by the set of the two operations R_{2y}, σ_{xz}.

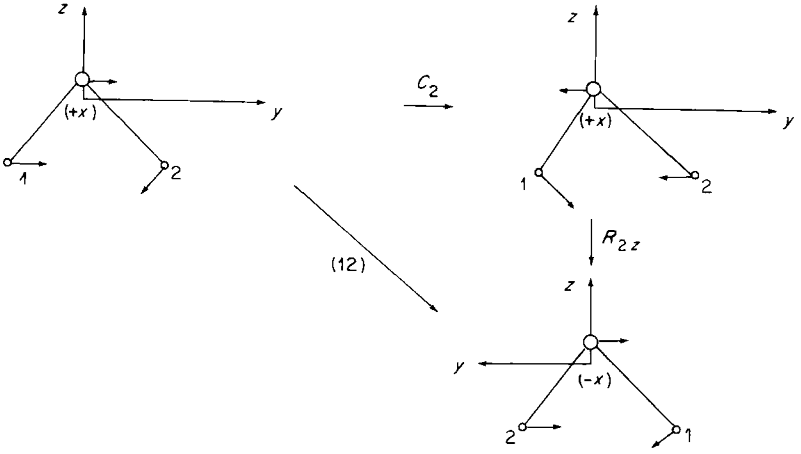


Fig. 4. PI group opertation on water molecule. The large circles represent the oxygen atom and the small ones the hydrogen atoms. The arrows represent their motions and the molecule is supposed to be taken in its equilibrium configuration.

Bunker and Papoušek (1969) have considered the particular case of the linear rigid molecules such as CO_2. For a molecule with $D_{\infty h}$ point group symmetry, the PI group consists of four elements E, (p), E^*, and (p^*); the operation (p) interchanges the coordinates of all pairs of identical nuclei which, in the equilibrium configuration, are located symmetrically about the center of mass of the molecule. For a molecule with $C_{\infty v}$ point group symmetry, the PI group consists of two elements E and E^*.

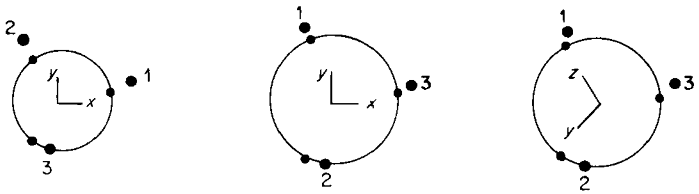
Thus, contrary to the case of nonlinear rigid molecules, there is not an isomorphism between the PI group and the point group for a linear rigid molecule. Bunker and Papoušek have shown that the operations of the PI group have a well-defined effect on the rovibronic coordinates and that the elements of the PI group leave the rovibronic Hamiltonian invariant. It is possible, therefore, to classify the rovibronic wave functions in the PI group. The symmetry labels $+s$, $-s$, $+a$, and $-a$ ordinarily used for the rovibronic levels of a $D_{\infty h}$ molecule are the labels of the four irreducible representations of the PI group (E, (p), E^*, $(p)^*$); the labels + and − used for the rovibronic levels of a $C_{\infty v}$ molecule are the labels of the irreducible representation of the PI group (E, E^*).

The elements of the point group in this case as in the case of the water molecule do not leave the rovibronic Hamiltonian invariant. Bunker and Papoušek have introduced a new group, the extended permutation inversion (EPI) group. The elements of the EPI group for a molecule with $D_{\infty h}$ point group symmetry are: E_ε, $(p)_\varepsilon$, E_ε^*, and $(p)_\varepsilon^*$, where the subscripts ε can assume any of the values in the range $0 \leqslant \varepsilon < 2\pi$ independently of each other. A general element $(p)_\varepsilon$ is defined by its effect on the space coordinates of the particles in the molecule and by its effect on the Euler angles θ, φ, and χ. It is by the effect on this last angle that a $(p)_\varepsilon$ operation differs from a p operation; the effect of a p operation on χ is undefined. Bunker and Papoušek have found that each element of the EPI group is the product of a rotation operation and an operation of the point group of the molecule and that the EPI group and the point group of a linear molecule are isomorphic. The effect of a general element P_ε of the EPI group of a linear molecule on a rovibronic wavefunction is the same as the effect of the element P of the PI group. The elements of the EPI group leave the rovibronic Hamiltonian invariant, and it is therefore possible to classify the rovibronic wavefunctions according to the representations of the EPI group of dimension one.

B. Nonrigid Molecules

Following Longuet-Higgins (1963) a molecule is said to be nonrigid if it is in an electronic state that has several potential energy minima separated by surmountable energy barriers. A typical example of a nonrigid molecule is the boron–trimethyl, $B(CH_3)_3$, in which the frame BC_3 is planar and in which the three CH_3 groups can rotate freely. For studying this kind of molecule Longuet-Higgins (1963) have introduced the concept of a molecular symmetry group or permutation inversion group (PI group). As seen before, this group consists of all feasible permutations P of the space and spin coordinates of identical nuclei in the molecule (including the identity E) and all feasible permutation inversions $P^* = PE^*$ (where E^* is the operation of inversion, in the molecular center of mass, of the space coordinates of all particles, nuclei, and electrons in the molecule). The rovibronic Hamiltonian is invariant by the permutations of the space and spin coordinates of identical nuclei and the inversion in the molecular center of mass.

Altmann (1971) has given several explanations concerning some difficulties of Longuet-Higgins method. He has shown that the permutation inversion operations used in this method are not true permutation inversion operations if the latter are defined as operations in which the axes (either laboratory or molecular) are kept fixed. In Fig. 5a, a true permutation (123) is shown in which the axes are not moved. In Fig. 5b, a Longuet-Higgins permutation L-H(123) is shown in which the molecular axes follow the nuclei.



a (123) permutation b (123) L-H permutation

Fig. 5. True and Longuet-Higgins permutations. The axes shown are the molecular axes.

It must be kept in mind that the E^* operation is not identical to the normal inversion i; whereas i inverts everything except the axes, E^* inverts the coordinates of all points as well as some of the molecular axes. The positioning of the axes after E^* must always be determined to preserve the constants of the motion as we have seen above for the rovibronic interactions. In fact, the choice of molecular axes is not so simple in the case of nonrigid molecules as in the case of rigid molecules; in a general asymmetric-top molecule, for example, two sets of Cartesian axes can be used, the instantaneous principal axis system and the internal motion axis system (Meyer and Günthard, 1968).

Altmann (1967, 1971) has studied nonrigid molecules from a different point of view. He has shown that the rovibronic wavefunctions of a nonrigid molecule can be classified in a group called the Schrödinger supergroup S. This group S is given in terms of two subgroups I (isodynamic) and G (point group) by the semidirect product (see Appendix 1)

$$S = I \rightsquigarrow G$$

where the subgroup I is the invariant subgroup.

The point group operations, the operations of G, are standard point group operations and transformations of axes for a given conformation of a nonrigid molecule. In such a molecule there are also operations which are not changes of axes, as for instance a displacement of atoms from one position of equilibrium to another. These operations are the isodynamic operations. The isodynamic group I is formed by the isodynamic generators, i.e., the isodynamic operations that are not themselves a product of an isodynamic operation times an operation of G.

If G is an invariant subgroup of S, then S is the direct product of I and G:

$$S = I \times G,$$

if G is the identity, $S = I$.

Altmann (1967, 1971) has given two different definitions of the isodynamic rotation in his two papers (Altmann, 1967, 1971) that are equivalent. Let us consider the rotation of a methyl group in boron trimethyl. In the isodynamic rotation of $+2\pi/3$, $(C_3{}^+)^{\mathrm{I}}$, each nucleus turns by $2\pi/3$ dragging its vibrational vector (Fig. 6) and the electronic function. After this operation a permutation $P_{(C_3{}^-)}$ relabels the nuclei in the same manner as the point group operation $C_3{}^-$. But, as we have seen earlier, in order to preserve the rovibronic Hamiltonian this isodynamic rotation has to be supplemented with a Euclidean operation. In such a case the latter operation, which we will call R^-, is a rotation of $-2\pi/3$ around the BC bond and the following relation can be written:

$$(132)_{\mathrm{LH\ perm}} = R^{-2\pi/3}(C_3{}^+)^{\mathrm{I}}$$

where LH perm means that the permutation (132) must be operated as it has been defined in Longuet-Higgins' theory.

Altmann also defined an isodynamic reflection through the BC_3 plane, called V^{I}; this operation is identical with (23)(56)(89)* in Longuet-Higgins' theory. Let us now consider the boron trimethyl molecule and let us call τ, τ', and τ'' the different torsional angles of the methyl groups

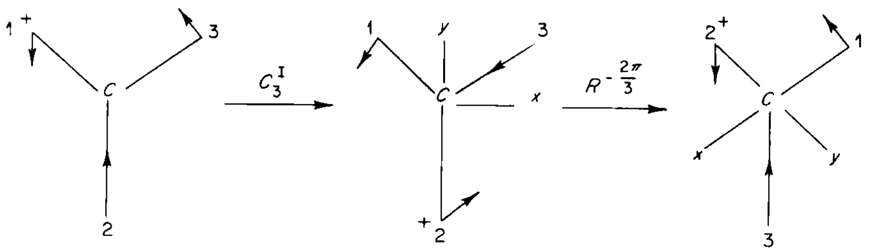


Fig. 6. Isodynamic rotation of a methyl group.

(τ is the angle between an X axis perpendicular to the BC bond and fixed in the BC_3 plane to a ξ axis attached to the CH_3 group). If the torsional angles τ, τ', and τ'' are unrelated, we call the structure a random structure; Altmann (1967) has shown that for the random structure of this molecule the supergroup S_r is:

$$S_r = [(\mathscr{C}_3{}^{\mathrm{I}} \times \mathscr{C}_3'^{\mathrm{I}} \times \mathscr{C}_3''^{\mathrm{I}}) \mathbin{\underset{\cdot}{\sim}} (C_3{}^{\mathrm{I}} \times V^{\mathrm{I}})] \mathbin{\underset{\cdot}{\sim}} W^{\mathrm{I}}$$

where $\mathscr{C}_3{}^{\mathrm{I}}$ is the isodynamic group and contains the identity and the two isodynamic rotations of $\pm 2\pi/3$ about the BC bond for one methyl group; $\mathscr{C}_3'^{\mathrm{I}}$ and $\mathscr{C}_3''^{\mathrm{I}}$ are the isodynamic groups for the two other methyl groups. The isodynamic group $\mathscr{C}_3{}^{\mathrm{I}}$ contains E, $C_3{}^{\mathrm{I}}$, and $(C_3{}^{\mathrm{I}})^-$. The operation $C_3{}^{\mathrm{I}}$ is a simultaneous rotation of the three methyl groups around their own axes such that τ, τ', τ'' are replaced by τ'', τ, τ'; in the other operation $(C_3{}^{\mathrm{I}})^-$, τ, τ', τ'' are replaced by τ', τ'', τ. The isodynamic group W^{I} contains the identity E and the exchange switch W^{I}. In an exchange switch, τ is changed into its negative and τ' and τ'' are first switched and then exchanged. This operation is identical with $(23)(475968)(bc)^*$ in Longuet-Higgins' theory.

There were some discrepancies between Altmann and Longuet-Higgins about the groups to use for boron–trimethyl. Longuet-Higgins (1963) found a group of 324 elements and Altmann (1967) proposed several groups depending on the considered configuration of $B(CH_3)_3$. In any case it is difficult to understand why the nature of the group depends on the studied configuration. After Watson (1971) pointed out that probably the definition of the isodynamic operation called $Z_2{}^{\mathrm{I}}$ was erroneous, Altmann (1971) explained that the flip $Z_2{}^{\mathrm{I}}$ is an acceptable operation if the interactions between the CH bonds and the BC_3 frame are neglected. If these interactions are taken into account, the correct isodynamic operation to be used is the one called V^{I} and there is no more disagreement between Altmann's and Longuet-Higgins' theories.

Finally, Gilles and Philippot (1972) recently analyzed the different approaches used from a formal point of view. By using the same development as Hougen, they showed that it is possible to write the nuclear permutations as products of operators acting on the BO variables. These operators are a proper rotation acting on the Euler angles and rotations and permutations respectively acting on the mobile coordinates and on the nuclear labels. This approach is similar to the one explained by Bunker and Papoušek (1969) in the case of linear molecules.

On the other hand, they (Gilles and Philippot, 1972) also defined isodynamic operators that describe the real motion of the molecule in the isodynamic operation. They also showed that these operators do not always form a group and that they cannot be used for classifying the rovibronic states. Due to this, they introduced other operators, the kinetic perrotations. This relates the final to the initial equilibrium nuclear configuration in the isodynamic transformation through a rotation in the mobile system coupled to a nuclear permutation. This approach is apparently similar to the one given by Altmann (1971).

III. Applications

The different theoretical points of view have been widely applied in different branches of spectroscopy. In each case the application of group theory to nonrigid molecules helps in the labeling of the states, in the calculation of the statistical weights, and more generally in the spectrum analysis. Some precise examples are given in microwave, infrared, ultraviolet, and NMR spectroscopy.

A. Microwave Spectra

The applications of microwave spectroscopy to molecules containing two or more symmetrically equivalent internal rotors are mainly due to works prior to Longuet-Higgins' theory. In a work on acetone (Myers and Wilson, 1960), it was shown that it is convenient to transform the Hamiltonian for overall and internal rotation written in the usual form (Swalen and Costain, 1959; Dreizler, 1961a,b) by introducing the new rotation coordinates $\alpha_+ = \frac{1}{2}(\alpha_1 + \alpha_2)$ and $\alpha_- = \frac{1}{2}(\alpha_1 - \alpha_2)$, and the new corresponding momenta $p_+ = (p_1 + p_2)$ and $p_- = (p_1 - p_2)$ where p_1 and p_2 are momenta associated with the internal rotation angles α_1 and α_2 (Fig. 7).

If we call C_{2x}, C_{2y} and C_{2z} the rotation of π about the axis x, y, and z, C_{31}^n the rotation of $2\pi n/3$ of only the methyl group 1 about its rotation

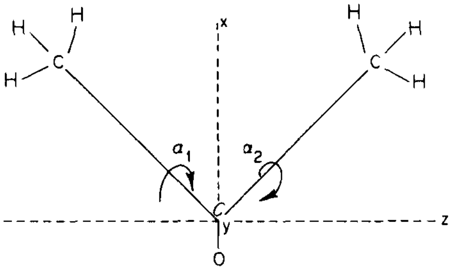


Fig. 7. Internal rotors in the acetone molecule.

axis, and C_{32}^{m} the rotation of $2\pi m/3$ of the methyl group 2, it can be checked that the operations

$$E, (C_{31}C_{32}^{2}, C_{31}^{2}C_{32}), (C_{2x}, C_{2x}C_{31}C_{32}^{2}, C_{2x}C_{31}^{2}C_{32})$$

leave α_+ unchanged. These operations form a group C_{3v}^{-} isomorphic to C_{3v} and they operate on α_-. Also the operations E, $(C_{31}C_{32}, C_{31}^{2}C_{32}^{2})$, $(C_{2z}, C_{2z}C_{31}C_{32}, C_{2z}C_{31}^{2}C_{32}^{2})$ leave α_- unchanged. These operations form a group C_{3v}^{+} isomorphic to C_{3v} and they operate on α_+. In consequence the operations that leave the total Hamiltonian invariant form a group I; this group is the direct product of C_{3v}^{+} by C_{3v}^{-}.

Since there are three irreducible representations in C_{3v} (two of dimension one, A_1 and A_2, and one of dimension two, E), there are nine irreducible representations in I (four of dimension one, A_1A_1, A_1A_2, A_2A_1, and A_2A_2, four of dimension two, A_1E, A_2E, EA_1, and EA_2, and one of dimension four, EE). In consequence nondegenerate levels, doubly degenerate levels, and quadruply degenerate levels are possible.

In acetone the dipole moment μ is along $0x$; this axis is in A_1 of C_{3v}^{-} and in A_2 of C_{3v}^{+} and μ is in A_1A_2. Consequently the following transitions are allowed:

$$A_1A_1 \leftrightarrow A_1A_2; \qquad A_2A_1 \leftrightarrow A_2A_2; \qquad A_1E \leftrightarrow A_1E;$$
$$A_2E \leftrightarrow A_2E; \qquad EA_1 \leftrightarrow EA_2; \qquad EE \leftrightarrow EE.$$

If the coupling between internal and overall rotation is neglected, approximate functions can be written in the form of an asymmetric rigid rotor function and an internal torsion function because the total Hamiltonian can be separated into a pure rotation term $H_R(\varphi, \theta, \chi)$, where φ, θ, χ are the Eulerian angles, and two pure torsion terms, $H_T(\alpha_+)$ and $H_T(\alpha_-)$. The coupling terms can be included by a perturbation calculation where these approximate functions are taken as zero order functions.

The asymmetric rigid rotor belongs to the group D_2 (Allen and Cross, 1963) and this group is a subgroup of the group I with the correlation between the representations of the two groups given in Table I.

TABLE I

Correlation of Representations Between Group I and Group D_2

Group I	A_1A_1	A_1A_2	A_2A_2	A_2A_1
Group D_2	A	B_x	B_y	B_z

In the high barrier case the wavefunctions for internal torsion can be approximated by the products of harmonic oscillator functions transformed by the group C_{3v}. If the torsion quantum number v is even or odd, the corresponding irreducible representations are $A_1 + E$ or $A_2 + E$. The possible symmetry species for the torsion product functions are given in Table II.

TABLE II

POSSIBLE SYMMETRY SPECIES FOR TORSION PRODUCT FUNCTIONS[a]

v_1	v_2	Symmetry species
e	e	$A_1A_1 + A_1E + EA_1 + EE$
o	o	$A_2A_2 + A_2E + EA_2 + EE$
e	o	$A_1A_2 + A_1E + EA_2 + EE$
o	e	$A_2A_1 + A_2E + EA_1 + EE$

[a] e = even, o = odd.

The symmetry species of the approximate total wavefunctions (Table III) are obtained from the correlation table (Table I) and from the multiplication of the C_{3v} irreducible representations. For the rigid rotor the allowed transitions are for the dipole moment along $0x$:

$$A \leftrightarrow B_x \qquad \text{and} \qquad B_z \leftrightarrow B_y$$

TABLE III

SYMMETRY SPECIES OF APPROXIMATED TOTAL WAVEFUNCTIONS

		Asymmetric rotor functions			
		A	B_x	B_z	B_y
Internal torsion functions	A_1A_1	A_1A_1	A_1A_2	A_2A_1	A_2A_2
	A_1A_2	A_1A_2	A_1A_1	A_2A_2	A_2A_1
	A_1E	A_1E	A_1E	A_2E	A_2E
	A_2A_1	A_2A_1	A_2A_2	A_1A_1	A_1A_2
	A_2A_2	A_2A_2	A_2A_1	A_1A_2	A_1A_1
	A_2E	A_2E	A_2E	A_1E	A_1E
	EA_1	EA_1	EA_2	EA_1	EA_2
	EA_2	EA_2	EA_1	EA_2	EA_1
	EE	EE	EE	EE	EE

If we take into account the selection rules found for the group I, we see that some transitions appear that are forbidden by considering only the rigid rotor function. For example, the transitions $EA_1 \leftrightarrow EA_2$ are allowed by the group I but the transitions $A \leftrightarrow B_y$ or $B_z \leftrightarrow B_x$ are forbidden for the rigid rotor and we have the following diagram:

$$\begin{tikzcd}
A \times EA_1 = EA_1 \arrow[r, leftrightarrow] \arrow[dr, leftrightarrow, dashed] & EA_2 = EA_1 \times B_x \\
B_z \times EA_1 = EA_1 \arrow[r, leftrightarrow] \arrow[ur, leftrightarrow, dashed] & EA_2 = EA_1 \times B_y
\end{tikzcd}$$

where the dotted lines exhibit the forbidden transitions that appear in addition to the allowed transitions.

These selection rules, established for the unperturbed Hamiltonian, are still valid for the total Hamiltonian after the rotation–torsion coupling terms and the torsion–torsion coupling terms have been introduced, indeed since the group I of the total Hamiltonian has been used, the symmetry and the number of the energy levels do not change and it is only the energy level positions that change. The introduction of these coupling terms is obtained by a perturbation calculation either in the special case of the torsional ground state or in the general case. From the splitting of the torsional ground state, the height of the potential barrier of dimethylallene (3-methyl-1,2-butadiene) has been found equal to 2025 cal/mole (Demaison, 1972; Demaison and Rudolph, 1971); this value is to be compared with the one for methylallene, 1589 cal/mole (Lide and Mann, 1957) and the one for methylketene, 1177 cal/mole (Bak *et al.*, 1966).

The microwave spectrum of *cis*-2-butene has been also studied in detail (Kondo *et al.*, 1970). It has been shown that the group under which the total Hamiltonian is invariant can be written in terms of Longuet-Higgins operations and is isomorphic to the group $C_{3v} \times C_{3v}$. By an analysis similar to the one carried out for dimethylallene, the barrier to internal rotation was found to be equal to 747 cal/mole. In microwave spectroscopy, these notions of symmetry are also useful for determining the statistical weights. Let us take the example of CH_3BF_2 (Longuet-Higgins, 1963) (Fig. 8). The protons are labeled 1, 2, 3 and the fluorines 4, 5.

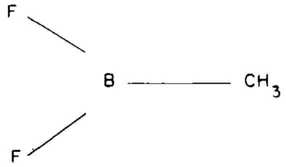


Fig. 8. The CH_3BF_2 molecule.

The molecular symmetry group of CH_3BF_2 is isomorphic to D_{3h} and the character table is given in Table IV. According to the exclusion principle, the overall wavefunction must belong to one or the other of the irreducible representations A_1'' or A_2''.

TABLE IV

CHARACTER TABLE OF MOLECULAR SYMMETRY GROUP OF CH_3BF_2

	E	(123) (132)	(23)* (31)* (12)*	(45)	(123)(45) (132)(45)	(23)(45)* (31)(45)* (12)(45)*
A_1'	1	1	1	1	1	1
A_2'	1	1	−1	1	1	−1
E'	2	−1	0	2	−1	0
A_1''	1	1	1	−1	−1	−1
A_2''	1	1	−1	−1	−1	1
E''	2	−1	0	−2	1	0

With the help of the character table it is easy to find the symmetry transformations of the nuclear spin states. For the protons, the possible spin states are $^4A_1'$ and $^2E'$; for the fluorine nuclei, the possible states are $^3A_1'$ and $^1A_1''$. According to Hougen (1962), it is possible to find how the wavefunctions of a symmetric top are transformed under rotations; the mass distribution in CH_3BF_2 is such that the molecule is nearly an oblate symmetric top. In such a case, the rotational species are A_1', A_1'', A_2', and A_2''. In conclusion, if the electronic–vibrational–torsional function is supposed to be totally symmetric, the symmetry of the overall function can be found by the product of the symmetries of the proton spin species, of the fluorine spin species, and of the rotational species. Since the overall function symmetry must be A_1'' or A_2'', the statistical weights of the pure rotational states of CH_3BF_2 are A_1' (4), A_1'' (12), A_2' (4), and A_2'' (12).

In the same way, the microwave spectra of CH_3SiH_3 and CH_3SiD_3 have been observed recently in the first excited state of the silyl rocking vibration ν_{12} (Hirota, 1972). The molecular symmetry group of methylsilane is G_{18} (Hougen, 1966; Bunker, 1965a), and the statistical weights determined with this symmetry group have been used in assigning the spectra.

B. Vibrational Spectra

In the case of nonrigid molecules, the vibrational spectra analysis must be completed with the molecular symmetry group. But if the experimental work is performed with a low resolution, the spectrum analysis may be carried out with a smaller group.

Let us first consider this approximate case. When the molecule can be divided into a rigid frame and a number of groups capable of internal rotation, Woodman (1970) has shown that the molecular symmetry group G is the semidirect product (symbolized by $\rtimes$) of a torsional subgroup H and a frame group F:

$$G = H \rtimes F.$$

If the molecule contains m equivalent groups, each containing n nuclei undergoing internal rotation, the subgroup H is an m-fold direct product of cyclic permutation groups $\mathscr{C}_n$ of order n:

$$H = (\mathscr{C}_n)^m.$$

The frame group contains all the permutations of the frame necessary to generate all the elements of the molecular symmetry group G. In the case of boron trimethyl $B(CH_3)_3$, the subgroup F is isomorphic with D_{3h} and

$$G = (\mathscr{C}_3)^3 \rtimes D_{3h}^F.$$

Starting with symmetry coordinates which transform as the irreducible representations of F, it is possible to build up the usual G and F matrices (Wilson *et al.*, 1955) and to diagonalize their product GF to obtain the normal vibrations. The kinetic energy matrix G is independent of the torsional angles but the force constant matrix F is not and contains off-diagonal elements. To overcome this difficulty it is possible to define a torsionally averaged force-constant matrix, $\bar{F}$, and to diagonalize $G\bar{F}$ and obtain torsionally averaged normal vibrations (Woodman, 1970). These normal vibrations transform as the irreducible representations of the frame group F.

Except in the case of high-resolution analysis, the usual practice has been to classify the normal vibrations according to the irreducible representations of the frame group. For example, in the case of ethyl- and n-propyl-benzene, the spectrum analysis has been performed with the C_{2v} group, the frame group if the phenyl group is defined as the frame and the alkyl group as the rotor (Saunders *et al.*, 1968; Harris and Thorley, 1972).

The vibrational analysis of the spectrum of molecules containing methyl groups such as dimethyl ether, dimethyl sulfide, dimethyl selenide, and dimethyl telluride had also been done with the corresponding frame group, C_{2v} (Freeman and Henshall, 1967). In the case of the oxides $(CH_3)_3NO$, $(CH_3)_3PO$, and $(CH_3)_3AsO$, it was also possible to describe all the vibrational modes with the exception of the torsional ones by using the frame group C_{3v} (Choplin and Kaufmann, 1970). It was found that in the case of $(CH_3)_3NO$ there are numerous couplings between the different frame vibrational modes and even between these vibrations and those of the group CH_3. The latter couplings forbid the use of the valence-force field (Wilson *et al.*, 1955) and a modified valence-force field with extradiagonal terms had to be used.

Starting with the molecular symmetry group written in Altmann's form, Bribes (1971) arrived at the same group as Woodman (1970) in the case of $B(CH_3)_3$. He has shown that in the case where only the vibrations are taken into consideration, the switches (Altmann, 1967) are the only operations to introduce in the isodynamic group. In an isodynamic rotation the vibrational Hamiltonian is kept invariant since the nuclear motion vectors are exchanged without any transformation of the coordinates. Therefore Bribes (1971) writes the vibrational group S_v in the following form:

$$S_v = I_v \underset{\cdot}{\sim} G_v,$$

where I_v is the isodynamic group including only the switches and G_v is the Wigner symmetry group of the chosen instantaneous configuration. In a switch U^{I}, there is a rotational motion of one part of the molecule, and the molecule goes from a given configuration to an enantiomorphic configuration. The result of a switch on the molecule CH_3NO_2 is given in Fig. 9.

The groups S_v for several molecules are given in Table V (Bribes, 1971).

This decomposition of the vibrational symmetry group has been applied to the spectrum analysis of $(CH_3)_2Hg$ (Bribes and Gaufrès, 1970),

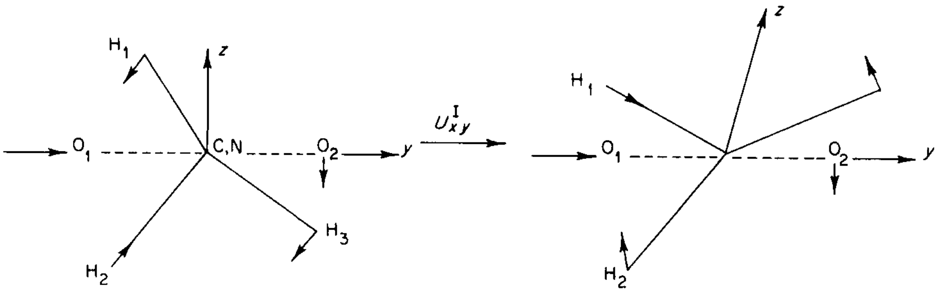


Fig. 9. The result of a switch on the molecule CH_3NO_2. The arrows represent the motions of the nuclei.

TABLE V

GROUPS S_v FOR SEVERAL MOLECULES[a]

Nonrigid molecules	Schrödinger supergroup S	S_v
CH_3NO_2	$(C_3^{\mathrm{I}} \times C_2^{\mathrm{I}}) \underset{\cdot}{\sim} U^{\mathrm{I}}$	$U^{\mathrm{I}} \equiv (E, U)$ isomorphic to Cs
$(CH_3)_2Hg$	$(C_3^{\mathrm{I}} \underset{\cdot}{\sim} U^{\mathrm{I}}) \underset{\cdot}{\sim} D_3$	$U^{\mathrm{I}} \underset{\cdot}{\sim} D_3$ isomorphic to D_{3h}
$B(CH_3)_3$	$(C_3^{\mathrm{I}} \times C_3^{\mathrm{I}} \times C_3^{\mathrm{I}}) \underset{\cdot}{\sim} (C_3^{\mathrm{I}} \underset{\cdot}{\sim} V^{\mathrm{I}}) \underset{\cdot}{\sim} W^{\mathrm{I}}$	$(C_3 \underset{\cdot}{\sim} V^{\mathrm{I}}) \underset{\cdot}{\sim} W^{\mathrm{I}}$ isomorphic to D_{3h}
$(C_2H_5)_2Hg$	$U^{\mathrm{I}} \underset{\cdot}{\sim} C_2$	$U^{\mathrm{I}} \underset{\cdot}{\sim} C_2$ isomorphic to D_2

[a] U^{I} = switch, V^{I} and W^{I} are defined in the theoretical part of this paper.

$(C_2H_5)_2Hg$ (Bribes and Gaufrès, 1971a; Guillermet *et al.*, 1971), and $C_2H_5NO_2$ (Guillermet *et al.*, 1971). This point of view may be also applied to the matrices F and G. Since the kinetic energy matrix F does not depend on the torsion angle but the force constant matrix G does depend on it, it is only possible to compute average force constants. This technique has been applied to dimethylacetylene (Duncan, 1964) and to dimethylmercury (Bribes and Gaufrès, 1971b).

In the case of a high resolution, the total molecular symmetry group must be used. One of the best examples of the high complication of the spectrum analysis in the case of molecules exhibiting internal rotation is given by the ethane-like molecules. The theory of the vibrational–rotational–torsional motions of ethane-like molecules with free internal rotation was discussed in detail in a recent series of papers (Bunker, 1965a,b 1967, 1968; Bunker and Hougen, 1967; Bunker and Longuet-Higgins, 1964; Hougen, 1964, 1965, 1966; Kirtman, 1964; Papoušek, 1968). Let us take the more precise example of dimethylacetylene, $CH_3 - C \equiv C - CH_3$ (Fig. 10).

The symmetry group to use for classifying the rotovibrational states of the molecule is the group introduced by Longuet-Higgins and called G_{36}. This group is the direct product of two groups of order 6, G_6' and G_6''. If we adopt the labeling of Fig. 10, the elements of G' are

$$\begin{array}{ll} E,\ (123)(465), & (14)(25)(36)(ab)(cd) \\ (132)(456), & (15)(26)(34)(ab)(cd) \\ & (16)(24)(35)(ab)(cd) \end{array}$$

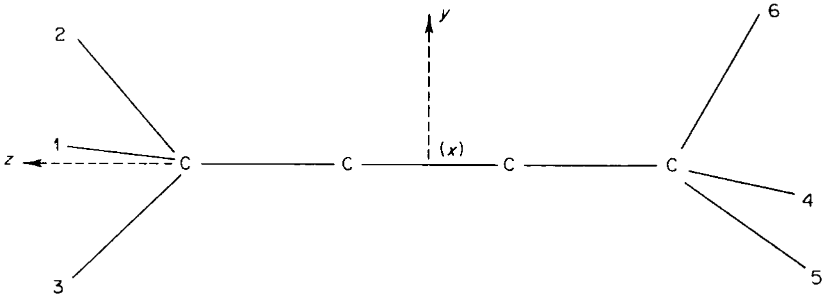


Fig. 10. Dimethylacetylene.

and those of G_6'' are:

$$\begin{array}{ll} E,\ (123)(456), & (14)(26)(35)(ab)(cd)^* \\ (132)(465), & (15)(24)(36)(ab)(cd)^* \\ & (16)(25)(34)(ab)(cd)^*. \end{array}$$

Since G_{36} is the direct product of two groups of order 6, the character table can be written at once. However, to determine the selection rules for the rotational, torsional, and vibrational parts of the wavefunctions separately, it was shown (Hougen, 1964) that it is necessary to use the double group of G_{36}, G_{36}^{+}.

If two rotor-fixed axis systems $x_t y_t z$ (top fixed) and $x_f y_f z$ (frame fixed) are defined (Fig. 11), the orientations of the $x_t y_t z$ and $x_f y_f z$ axes relative to the space-fixed XYZ axes are measured by the Eulerian angles (θ, φ, χ_t) and (θ, φ, χ_f), respectively. If the angles $\chi = (\chi_t + \chi_f)/2$ and $\gamma = (\chi_t - \chi_f)/2$ are introduced, the rotovibrational wavefunctions of dimethylacetylene can be written as the product of three functions, a rotational function of the Eulerian angles θ, φ, and χ, a torsional function of the angle γ, and a vibrational function which may involve γ. When the molecule is in a particular configuration, χ can have the value χ or $(\chi + \pi)$ and γ the value γ or $(\gamma + \pi)$. The angles χ and γ are double-valued functions of molecular configurations. Therefore the rotational, torsional, and vibrational functions can be double valued and instead of the group G_{36} the double group G_{36}^{+} must be used in particular to determine infrared and Raman selection rules for the rotational, torsional, and vibrational quantum numbers.

The main problem with a molecule such as dimethylacetylene is that many of the symmetry coordinates involve γ as a parameter and since γ varies with time, the G matrix calculation cannot be the same as for a rigid molecule. Usually (Bunker, 1967) the symmetry coordinates are chosen so

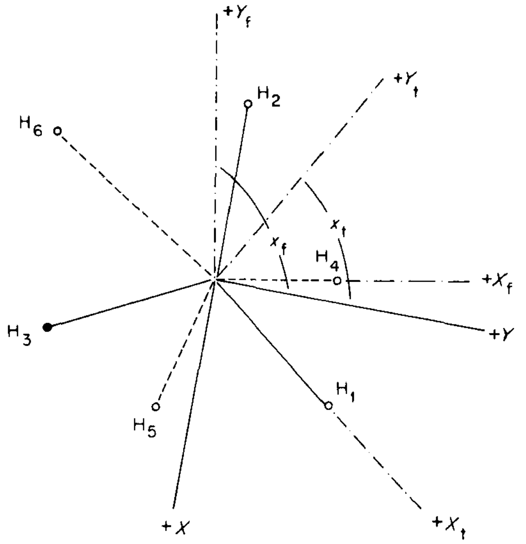


Fig. 11. Two rotor-fixed axis systems, where the $x_t y_t z$ system is top fixed and the $x_f y_f z$ system is frame fixed.

that the G and F matrices are block diagonal to the maximum extent and so that the G matrix is independent of γ. But the elements of the F matrix depend on γ. In such a molecule, the FG calculation has to be performed for several values of γ. The infrared and Raman selection rules can be established (Hougen, 1964; Bunker, 1965a,b) in a way parallel to that used for the rigid molecules.

The high-resolution infrared spectrum of dimethylacetylene in the 3000 cm^{-1} region was analyzed recently (Olson and Papoušek, 1971) and the double group G_{36}^{+} was used for the classification of the states of dimethylacetylene. The upper limit to the torsional barrier was estimated to be 4 cm^{-1} (11.4 cal) from the partial resolution of the Q branches.

The symmetry coordinates and the factorization of the FG matrices were also discussed for ethylene-like molecules X_2Y_4 (Papoušek *et al.*, 1971) and nitromethane-like molecules (Papoušek *et al.*, 1971; Fleming and Banwell, 1969). In the case of X_2Y_4, the group to use is G_{16}^{+} but the kinematic matrix G does not factorize completely according to the irreducible species of the symmetry coordinates. There are interaction terms depending on the angle of internal rotation relating the orientation of the top half to the orientation of the frame half of the molecule.

In toluene, $C_6H_5CH_3$, a nitromethane-like molecule, the rotation barrier is very small (13.9 cal/mole) as determined by microwave spectroscopy (Rudolph *et al.*, 1967). The Longuet-Higgins group of toluene (Turrell, 1970) is the same as the one of CH_3–BF_2. If the atoms of the toluene molecule are labeled as shown in Fig. 12, and if the atoms 8–11 are represented by a and the atoms 12–15 by b, the group is formed by the following twelve operations: E, (123), (132), (23)*, (31)*, (12)*, (ab), $(123)(ab)$, $(132)(ab)$, $(23)(ab)^*$, $(31)(ab)^*$, and $(12)(ab)^*$. This group is isomorphic with the D_{3h} group. The representation of the vibrational modes including the translational and rotational motions is determined by considering the effect of each symmetry operation on the sets of axes fixed on the atoms. After subtracting the representations for the translation and the rotation, the representation Γ_R of the vibrational modes is given by:

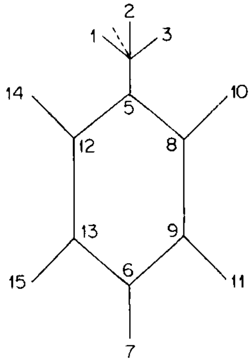


Fig. 12. Labeling of the atoms in the toluene molecule.

$$\Gamma_R = 13A_1' + 4A_2' + 13A_1'' + 9A_2''.$$

Since all the representations included in Γ_R are of dimension one, this confirms the experimental fact that the asymmetric stretching fundamentals of the methyl group remain clearly split over the temperature range −60°C to 200°C (Turrell, 1970).

Also with the help of the molecular symmetry group, the assignments have been made for the 36 normal frequencies of vibration of cyclopentanone (Howard-Lock and King, 1970), the infrared and Raman selection rules have been established for cyclopentane (Mills, 1971), and the antisymmetric wagging band of hydrazine has been analyzed (Hamada *et al.*, 1970).

C. Vibronic Spectra

The groups introduced in Section II of this review can help to analyze a vibronic absorption band. For example, the electronic bands of phenol vapor in the region 2500–2900 Å (Bist *et al.*, 1966, 1967a,b; Brand *et al.*, 1968) have been analyzed, and the vibrational structure associated with the in-plane modes has been discussed with the help of the molecular symmetry group. The ground state of phenol is planar. The excited state is supposed to be planar also, and the point group for both states is C_s and the electronic transition analogous to the $B_{2u} \leftarrow A_{1g}$ transition of benzene is $A' \leftarrow A'$. Phenol has equivalent conformations exchanged between each other by internal rotation through π about the CO bond. The existence of these two equivalent conformations results in a degeneracy which is split when tunneling is taken into account. The rovibronic and vibronic states of the isolated molecule can be classified by the species of a molecular symmetry group.

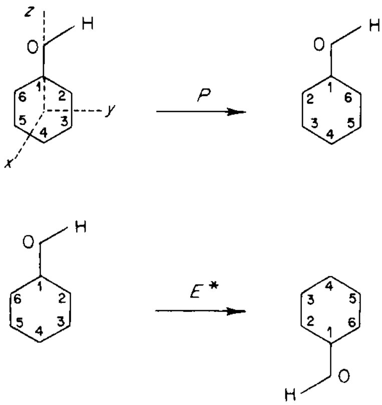


Fig. 13. The effect of the operations P and E^* on the phenol molecule.

Following Longuet-Higgins' (1963) notation, this group of order 4 includes the following elements: the identity E, the inversion E^*, the permutation P, and the operation $P^* = PE^* = P^*E$. The effect of the operations P and E^* is summarized in Fig. 13. The permutation P exchanges the C_2C_6, H_2H_6, C_3C_5, H_3H_5 pairs of nuclei and has the effect of converting one conformation into the other one across the barrier. The operation E^* leaves the molecule in its original geometrical conformation.

The character table for the molecular group G_4 is given in Table VI. For finding the effect of the operations on the axes z, x, and y and their

TABLE VI

CHARACTER TABLE FOR G_4

G_4	E	P	E^*	P^*	
A_1	1	1	1	1	z, x^2, y^2, z^2
A_2	1	1	-1	-1	xy
B_1	1	-1	-1	1	x, xz
B_2	1	-1	1	-1	y, yz

products, it must be remembered that the Hamiltonian is invariant under the rotations $C_2{}^z$ and $C_2{}^x$, and the operation P must be followed by the rotation $C_2{}^z$ for converting one conformation into the other. In the same manner the operation E^* must be followed by $C_2{}^x$.

In order to investigate the transformation properties of the internal coordinates, the torsional angle τ is defined so that the planar conformations have $\tau = \pm\frac{1}{2}\pi$ and the four operations of G_4 transform τ as follows:

$$E\colon \tau \to \tau, \qquad P\colon \tau \to \tau + \pi,$$

$$E^*\colon \tau \to -\tau + \pi, \qquad P^*\colon \tau \to -\tau.$$

If the torsional wavefunctions can be written as a sum or difference of harmonic oscillator wavefunctions,

$$2^{-1/2}[\Psi_v(\tau + \tfrac{1}{2}\pi) \pm \Psi_v(\tau - \tfrac{1}{2}\pi)],$$

the symmetry properties of the symmetric and antisymmetric combinations are given in Table VII. It can be shown (Bist *et al.*, 1966) that the representation Γ of the normal vibrations of the C_6H_5O fragment in the group G_4 is given by the relation

$$\Gamma = 11a_1 + 9b_2 + 3a_2 + 7b_1.$$

TABLE VII

SYMMETRY PROPERTIES OF SYMMETRIC AND ANTISYMMETRIC COMBINATIONS

Symmetric	v even	$2^{-1/2}[\Psi_v(\tau + \frac{1}{2}\pi) + \Psi_v(\tau - \frac{1}{2}\pi)]$	A_1
Symmetric	v odd	$2^{-1/2}[\Psi_v(\tau + \frac{1}{2}\pi) - \Psi_v(\tau - \frac{1}{2}\pi)]$	B_1
Antisymmetric	v even	$2^{-1/2}[\Psi_v(\tau + \frac{1}{2}\pi) - \Psi_v(\tau - \frac{1}{2}\pi)]$	B_2
Antisymmetric	v odd	$2^{-1/2}[\Psi_v(\tau + \frac{1}{2}\pi) + \Psi_v(\tau - \frac{1}{2}\pi)]$	A_2

The excited state electronic symmetry is B_2 in G_4. The four different vibrational symmetries are a_1, b_2, a_2, and b_1. For $v = 0^{\pm}$ (and all even-numbered torsional levels) we have the symmetries A_1 and B_2. The symmetries of the excited state sublevels are thus the following:

$$\begin{array}{ll} B_2 \otimes a_1 \otimes A_1 = B_2, & B_2 \otimes a_1 \otimes B_2 = A_1, \\ B_2 \otimes b_2 \otimes A_1 = A_1, & B_2 \otimes b_2 \otimes B_2 = B_2, \\ B_2 \otimes b_1 \otimes A_1 = A_2, & B_2 \otimes b_1 \otimes B_2 = B_1, \\ B_2 \otimes a_2 \otimes A_1 = B_1, & B_2 \otimes a_2 \otimes B_2 = A_2, \end{array}$$

where the lower case letters give the vibrational symmetry. The $0^{\pm}$ sublevels are $A_1 + B_2$ in the excited state for the vibrations a_1 and b_2 and also in the ground state of electronic symmetry A_1 and torsional symmetries A_1 and B_2. The G_4 selection rules allow y-polarized $+\leftrightarrow+$ and $-\leftrightarrow-$ and z-polarized $+\leftrightarrow-$ combinations between sublevels. By computing the leading term in the expression for the intensity of an allowed transition, it can be shown that the 0–0 band will be essentially y-polarized and will have a small z-polarized component. In fact, experimentally (Bist *et al.*, 1966) the spectrum comprises a strong allowed system of bands polarized in the plane of the molecule along $0y$ and a group of much weaker forbidden subsystems polarized along $0z$. All a_1 vibrational frequencies of the phenyl group in the excited state of phenol have been assigned. All these facts are summarized in Fig. 14. By the determination of the torsional infrared frequencies and of the subtorsional splittings by microwaves, the

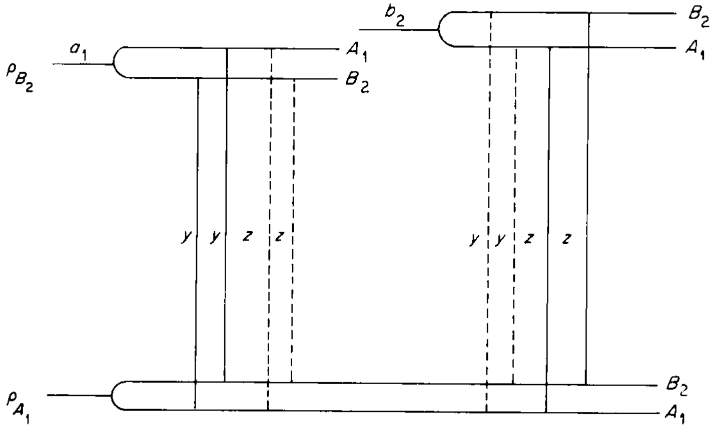


Fig. 14. The electronic transition $B_2 \leftarrow A_1$ in the phenol molecule.

barrier hindering internal rotation was estimated to be 1215 cm^{-1} (Bist *et al.*, 1967b).

In the same manner, a complete classification of the vibronic sublevels of aniline associated with the 2940 Å band system (Brand *et al.*, 1966, 1971) is accomplished by using a group G_8 of order eight. Twisting and inversion split a symmetric level into sublevels $A_g + B_{2g} + B_{1u} + B_{3u}$. If twisting and inversion are separately taken into consideration, it is found that inversion divides a totally symmetrical rovibronic state into A_g and B_{2g} components. Torsion divides the A_g state into A_g and B_{1u} components and the B_{2g} state into B_{2g} and B_{3u}. The inversion splittings are much greater than the torsional splittings. The excited aniline has the electronic species B_{3g} in G_8; for an a_g vibration, the sublevels are of B_{3g}, B_{2u}, B_{1g}, and A_u symmetry and the corresponding transitions are y-polarized (Fig. 15). The inversion doubling has been estimated as 41 ± 9 cm^{-1} (Brand *et al.*, 1971).

The vibrational and rotational analyses of the 3500 Å $^1A_2$–$^1A_1$ transition of formaldehyde were also carried out with the corresponding molecular symmetry (Job *et al.*, 1969). The operations applicable to H_2CO and D_2CO are the identity E, the permutation (x_1, x_2) where $x = H$ or D, the inversion E^* and the permutation inversion $(x_1, x_2)^*$. These four operations constitute a group (G_4) isomorphic with the C_{2v} point group.

In microwave, infrared, or ultraviolet spectroscopy these group considerations on nonrigid molecules are now widely applied; now we will see that these considerations applied in NMR can help to factorize the NMR Hamiltonian matrix.

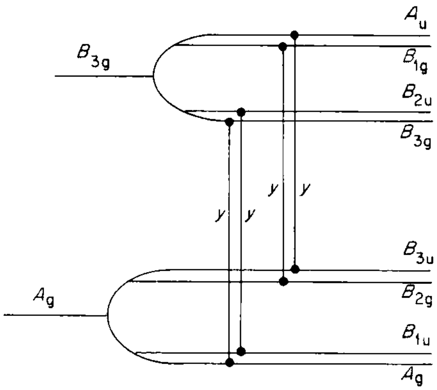


Fig. 15. The electronic transition $B_{3g} \leftarrow A_g$ in the aniline molecule.

D. NMR Spectra

The magnetic energy levels of a set of nuclei can be considered as the eigenvalues of a magnetic Hamiltonian

$$\mathscr{H} = \sum_{X} \nu_X I_z^{(x)} + \sum_{X<Y}\sum J_{XY} \mathbf{I}^{(x)} \cdot \mathbf{I}^{(y)},$$

where $I_z^{(x)}$ is the nuclear spin of the nucleus X and $\mathbf{I}^{(x)} \cdot \mathbf{I}^{(y)}$ the scalar product of the spin operators of each pair (X, Y) of nuclei; ν_X is the resonance frequency that depends on the magnetogyric ratio g_X of the nucleus X, on the applied field H_0, on the shielding constant σ_X and on the nuclear magneton β_N. There is the following relation:

$$\nu_X = g_X \beta_X (1 - \sigma_X) H_0 .$$

The coupling constant between nuclei X and Y, J_{XY}, depends on the surroundings of X and Y, for example, on the number and the nature of the chemical bonds existing between X and Y in the molecule.

Both ν_X and J_{XY} are parameters which are averaged over a period of the order of the nuclear spin relaxation times ($\sim 1 s$). These quantities are therefore averaged over all vibrorotational and torsional motions except if high barriers are involved during these motions.

The symmetry group adapted to nuclear magnetic resonance problems can be defined as containing all the permutations which do not change the parameters of the nuclei, such as coupling constants and chemical shifts (Corio, 1966). Nuclei that can be exchanged by such permutations are said to be equivalent; for them, the following conditions are fulfilled:

(1) The corresponding chemical shifts are equal: $\nu_i = \nu_j = \nu_k = \cdots$ for the nuclei $i, j, k, l, \ldots$.

(2) The coupling constants are equal to one another: $J_{ij} = J_{ik} = J_{il} = J_{kl} = \cdots$.

(3) The coupling constants of these nuclei to any other nucleus Z of the molecule are equal: $J_{iz} = J_{jz} = J_{kz} = \cdots$.

In consequence the elements of the NMR symmetry group for a molecule with one set of magnetically equivalent nuclei are the $n!$ permutations of the n equivalent nuclei, and the corresponding wavefunctions can be established with the projection operators connected with the different irreducible representations of the permutation group S_n on n symbols.

If a molecule contains two different sets N and N' of equivalent nuclei, the determination of the spin wavefunctions can be done with the help of the projection operators of the group G, which is the direct product of the groups S_n and $S_{n'}$ (permutation groups of n and n' symbols):

$$G = S_n \times S_{n'}.$$

But a more interesting case arises when the two sets N and N' of equivalent nuclei can be exchanged by a symmetry operation s. In such a case there is the following relation between S_n and $S_{n'}$:

$$sS_n s^{-1} = S_{n'},$$

and it can be said that the operation s conjugates S_n and $S_{n'}$. The appropriate NMR symmetry group G is the semidirect product (Serre, 1967) (represented by the symbol $\underset{\cdot}{\sim}$) of the invariant subgroup $S_n \times S_{n'}$ and a subgroup of two elements, 1 and s:

$$G = (S_n \times S_{n'}) \underset{\cdot}{\sim} (1, s).$$

In general there is more than one operation exchanging the two sets N and N' and the subgroup $(1, s)$ is not an invariant subgroup of the total groups.

For example, in the ethane molecule $^{13}CH_3$–$^{13}CH_3$, the rotation of the two CH_3 groups is much faster than the timing of an NMR experiment, and the NMR symmetry group G (Serre, 1967, 1969) is:

$$G = (D_3 \times D_3') \underset{\cdot}{\sim} (1, s).$$

Woodman (1970) has defined an effective NMR symmetry group (ENMRSG) which allows the expression of the concept of magnetic equivalence in group theory. A magnetically equivalent set of nuclei fulfills the conditions (1) and (3) of a set of completely equivalent nuclei but it does not fulfill the condition (2); the nuclei are not equally coupled to each other. It is well known (Corio, 1966) that the couplings in a magnetically equivalent group do not modify the spectrum. In such a case it is possible to define an effective NMR Hamiltonian where all the couplings in magnetically equivalent sets are put equal to zero; in consequence the ENMRSG contains all the possible permutations within a magnetically equivalent group.

For example, in the case of the molecule $B(^{12}CH_3)_3$, the NMR symmetry group G_{NMR} is:

$$G_{NMR} = (D_3 \times D_3' \times D_3'') \underset{\cdot}{\sim} K,$$

where the subgroup K contains all the symmetry operations that allow the exchange of one CH_3 group with another one; K is isomorphic to D_3 and as has been shown by Woodman (1970):

$$G_{\mathrm{NMR}} = (D_3 \times D_3' \times D_3'') \rtimes D_3 .$$

But if all the coupling constants J_{ij} between the hydrogens are taken equal to zero, and as the carbon nuclei are not magnetic, the nine hydrogens are magnetically equivalent and the effective NMR symmetry group G_{E} is:

$$G_{\mathrm{E}} = S_9 .$$

Factorization of the NMR Hamiltonian matrix is thereafter completed by using the spin wavefunctions determined with this group, and the matrix is factorized to the same extent as in the pure magnetic equivalence approach (Woodman, 1966).

Appendix 1: Semidirect Product

Let G be a group and A and B two subgroups of G. Then G is a semidirect product of B by A if the following conditions are satisfied:

(1) A is an invariant subgroup of G (for every $x \in G$, $xAx^{-1} = A$).

(2) Every element x, $x \in G$, can only be written in one manner under the form $a \cdot b$ with $a \in A$ and $b \in B$. Then we have $G = A \rtimes B$.

Appendix 2: Three Theorems

Three theorems are necessary to determine the irreducible representations of groups which are semidirect products of a subgroup by an invariant subgroup. Let us first recall the definitions of subspaces of representation and of an induced representation.

(a) Subspaces of representation. Let V be a vector space over the field of the complex numbers. Let $GL(V)$ be the group of the isomorphisms of V onto itself. Let G be a finite group. A linear representation of G is a homomorphism ρ of the group G in the group $GL(V)$.

Let the decomposition of V in direct sum of subspaces be $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$; we suppose that: (a) all the W_i are permuted by G, so that if $s \in G$, $\rho(s)W_i$ is one of the W_j; (b) G permutes the W_i transitively; in other words, for i and j given, there is one s, $s \in G$, such that $\rho(s)W_i = W_j$.

(b) Induced representation. Let W_1 be one of the subspaces; let H be the subgroup of G, where the elements s are such that

$$\rho(s)W_1 = W_1, \qquad s \in G,$$

W_1 is a representation of H. The representation of G induced by the representation of H in W_1 is V.

If the representation of H in W_1 and if the subgroup H of G are known, it is possible to reconstitute V and there is unicity.

The three theorems are the following:

Theorem 1. Character of an Induced Representation. *Let H be a subgroup of a group G, $H \subset G$, ρ be a representation of H of character χ, ρ^* be the representation of G induced by ρ, and χ^* be the character of ρ^*. Then*

$$\chi^*(g) = \frac{1}{h} \sum_{\substack{xgx^{-1} \in H \\ x \in G}} \chi(xgx^{-1})$$

where h is the order of the subgroup H. We see, particularly, that

$$\chi^*(g) = 0$$

if g is not conjugate with any element of H and

$$\chi^*(g) = 0$$

for every element g on the outside of H if H is an invariant subgroup.

Theorem 2. Corollary to Mackey's Irreducibility Criterion. *Let H be an invariant subgroup of a group G and ρ_s be the conjugate representation of a representation ρ,*

$$\rho_s(x) = \rho(s^{-1}xs) \qquad \textit{for} \quad x \in H \quad \textit{and} \quad s \notin H, \quad s \in G.$$

The necessary and sufficient condition for the induced representation $V = \mathrm{Ind}_H{}^G(\rho)$ (the representation V of G induced by the representation ρ of H) to be irreducible is that ρ is irreducible and is not isomorphic to any of its conjugates ρ_s for s in G and not in H.

Theorem 3. The Case of an Abelian Invariant Subgroup. *Let G be a group such that $G = H \rtimes X$ where H is an abelian invariant subgroup.*

Let χ be an irreducible representation of H and its character; it is of dimension one since H is Abelian.

Let X_χ be the set of x such that

$$\chi(xhx^{-1}) = \chi(h) \qquad \textit{for every } h \in H.$$

Let Ψ *be an irreducible representation of* X_x *and its character. Let* G_x *be the semidirect product of* H *and* X_x,

$$G_x = H \curvearrowright X_x .$$

G_x *is a subgroup of* G.

(1) *There is an irreducible representation* $r_{\chi,\Psi}$ *of* G *whose character is* $\chi\Psi$*: For an element* $h \,.\, x$, $h \in H$ *and* $x \in X_x$, *the character of this representation is*

$$\chi(h)\Psi(x).$$

(2) *Let* $r^*_{\chi,\Psi}$ *be the representation of* G *induced by the representation* $r_{\chi,\Psi}$ *of the subgroup of* G, G_X. *Then* $r^*_{\chi,\Psi}$ *is irreducible. Every irreducible representation of* G *can be obtained in such a way for proper* χ *and* Ψ.

Stone (1964) has described a method to calculate character tables for the symmetry groups of molecules consisting of methyl groups attached to a rigid framework. But in this method, it is necessary to know the basis functions. Several authors (Lamb, 1966; Serre, 1968; Woodman, 1970) have worked on this subject. In this paper we will describe the complete building up of the character table of the rather simple group K necessary for ethane in NMR:

$$K = (D_3 \times D_3') \curvearrowright (1, s).$$

In this group D_3 and D_3' are the permutation groups of the hydrogens of the two methyls and s is one of the symmetries exchanging the two methyls. The conjugation relation for this group is

$$sD_3s^{-1} = D_3'.$$

Let us call χ_1, χ_{-1}, and χ_2 the irreducible representations of D_3 (χ_1', χ'_{-1}, and χ_2' those of D_3'); the irreducible representations of $D_3 \times D_3'$ are either invariant by s ($\chi_i \otimes \chi_i'$ with $i = 1, -1, 2$) or noninvariant by s ($\chi_i \otimes \chi_j'$ with $i \neq j$ and i or $j = 1, -1, 2$). The induced representations in the group K of the representations invariant by s are reducible and those of the representations noninvariant by s are irreducible.

(a) Classes of the group K. Let C be a class of D_3 and C' a class of D_3'. A class of $D_3 \times D_3'$ is constituted by the pair (C, C') but in the group K the pair (C', C) is in the same class owing to the conjugation relation. The group ($D_3 \times D_3'$) has nine classes but the invariant subgroup ($D_3 \times D_3'$) of K brings $3 \times 4/2 = 6$ classes in K. If we call w and τ the rotation of $+2\pi/3$ and the symmetry in D_3, w' and τ' those in D_3', these six classes $A, B, \ldots, F$ are given in Table VIII.

TABLE VIII

CLASSES OF THE GROUP K

Classes	Operations
A	E
B	w, w^2, w', w'^2
C	$\tau, w\tau, w^2\tau, w', w'\tau', w'^2\tau'$
D	$ww', ww'^2, w^2w', w^2w'^2$
E	$w'\tau, w'w\tau, w'w^2\tau, w'^2\tau, w'^2w\tau, w'^2w^2\tau, w\tau', ww'\tau', ww'^2\tau', w^2w'\tau', w^2w'^2\tau', w^2\tau',$
F	$\tau\tau', \tau w'\tau', \tau w'^2\tau', w\tau\tau', w\tau w'\tau', w\tau w'^2\tau', w^2\tau\tau', w^2\tau w'\tau', w^2\tau w'^2\tau'$

At this point it remains to determine when two elements $g_1g_2's$ and $g_3g_4's$ are in the same class (g_1 and $g_3 \in D_3$, g_2' and $g_4' \in D_3'$). They are conjugate by an element of the form XY' ($X \in D_3$, $Y' \in D_3'$) if

$$g_1g_2's = X^{-1}Y'^{-1}g_3g_4'sY'X.$$

By applying the conjugation relation it can be shown that the necessary and sufficient condition for $g_1g_2's$ and $g_3g_4's$ to be conjugate is that the elements g_1g_2 and g_3g_4 are themselves conjugate in D_3. The conjugation by an element $XY's$ gives the same condition. Three additional classes A', B', C' are found in that manner and their elements are given in Table IX.

(b) Irreducible representations of the group K. The irreducible representations of $D_3 \times D_3'$ are given in Table X.

The conjugate representations χ_2'' and χ_4'' have the same induced representation in the group K and this one is irreducible. It is the representation Γ_5 given in Table XI. The situation is the same for the representations χ_3'' and χ_7'', χ_6'' and χ_8''. The induced representation by χ_3'' (or by χ_6'') is Γ_6 (or Γ_7).

TABLE IX

ADDITIONAL CLASSES OF THE GROUP K

Classes	Operations
A'	$s, ww'^2s, w^2w's, \tau\tau's, w\tau w'\tau's, w^2\tau w'^2\tau's$
B'	$ww's, w^2w'^2s, ws, w^2s, w's, w'^2s, \tau w'\tau's, \tau w'^2\tau's, w\tau\tau's, w^2\tau\tau's, w\tau w'^2\tau's, w^2\tau w'\tau'$
C'	$\tau s, w\tau s, w^2\tau s, \tau's, w'\tau's, w'^2\tau's, w'\tau, w'w\tau s, w'w^2\tau s, w\tau's, ww'\tau's, ww'^2\tau's, w'^2\tau, w'^2w\tau s, w'^2w^2\tau s, w^2\tau's, w^2w'\tau's, w^2w'^2\tau's$

TABLE X

Irreducible Representations of the Group $D_3 \times D_3'$

Classes	E	w	τ	w'	$w'w$	$w'\tau$	τ'	$\tau'w$	$\tau'\tau$
Number of elements	1	2	3	2	4	6	3	6	9
$\chi_1 \otimes \chi_1' = \chi_1''$	1	1	1	1	1	1	1	1	1
$\chi_1 \otimes \chi_{-1}' = \chi_2''$	1	1	1	1	1	1	−1	−1	−1
$\chi_1 \otimes \chi_2' = \chi_3''$	2	2	2	−1	−1	−1	0	0	0
$\chi_{-1} \otimes \chi_1' = \chi_4''$	1	1	−1	1	1	−1	1	1	−1
$\chi_{-1} \otimes \chi_{-1}' = \chi_5''$	1	1	−1	1	1	−1	−1	−1	1
$\chi_{-1} \otimes \chi_2' = \chi_6''$	2	2	−2	−1	−1	1	0	0	0
$\chi_2 \otimes \chi_1' = \chi_7''$	2	−1	0	2	−1	0	2	−1	0
$\chi_2 \otimes \chi_{-1}' = \chi_8''$	2	−1	0	2	−1	0	−2	1	0
$\chi_2 \otimes \chi_2' = \chi_9''$	4	−2	0	−2	1	0	0	0	0

TABLE XI

Irreducible Representations of the Group K

	A	B	C	D	E	F	A'	B'	C'
Γ_1	1	1	1	1	1	1	1	1	1
Γ_2	1	1	1	1	1	1	−1	−1	−1
Γ_3	1	1	−1	1	−1	1	1	1	−1
Γ_4	1	1	−1	1	−1	1	−1	−1	1
Γ_5	2	2	0	2	0	−2	0	0	0
Γ_6	4	1	2	−2	−1	0	0	0	0
Γ_7	4	1	−2	−2	1	0	0	0	0
Γ_8	4	−2	0	1	0	0	2	−1	0
Γ_9	4	−2	0	1	0	0	−2	1	0

At this point we must find the irreducible representations of K corresponding to those of $(D_3 \times D_3')$ invariant by s. Let us look at the representation χ_9'' for example. The subspace V of the representation χ_2 is of dimension two: Let us call e and e' the basis vectors. In $V \otimes V$ the basis is

$$e \otimes e, \qquad e \otimes e', \qquad e' \otimes e, \qquad e' \otimes e'.$$

By the action of s, we have

$$\begin{aligned} s(e \otimes e) &= e \otimes e &\quad \text{or} \quad& s(e \otimes e) = -(e \otimes e),\\ s(e \otimes e') &= e' \otimes e &\quad \text{or} \quad& s(e \otimes e') = -(e' \otimes e),\\ s(e' \otimes e) &= e \otimes e' &\quad \text{or} \quad& s(e' \otimes e) = -(e \otimes e'),\\ s(e' \otimes e') &= e' \otimes e' &\quad \text{or} \quad& s(e' \otimes e') = -(e' \otimes e'). \end{aligned}$$

The character associated with the element s is ± 2. An element $g_1 g_2'$ ($g_1 \in D_3$, $g_2' \in D_3'$) operates by $\chi_2 \otimes \chi_2'$; then for this element the character is the same as that in the representation χ_9''.

If we look for the character associated with an element $g_1 g_2' s$ we have

$$\begin{tikzcd}
e \otimes e \arrow[r, "s"] & e \otimes e \arrow[r, "g_1 g_2'"] & g_1 e \otimes g_2' e, \\
e \otimes e' \arrow[r, "s"] & e' \otimes e \arrow[r, "g_1 g_2'"] & g_1 e' \otimes g_2' e, \\
e' \otimes e \arrow[r, "s"] & e \otimes e' \arrow[r, "g_1 g_2'"] & g_1 e \otimes g_2' e', \\
e' \otimes e' \arrow[r, "s"] & e' \otimes e' \arrow[r, "g_1 g_2'"] & g_1 e' \otimes g_2' e'.
\end{tikzcd}$$

Let us put

$$g_1 e = ae + be', \qquad g_2' e = \alpha e + \beta e',$$
$$g_1 e' = ce + de', \qquad g_2' e' = \gamma e + \delta e'.$$

The trace of the matrix associated with the element $g_1 g_2' s$ is

$$\text{Trace} = a\alpha + d\delta + c\beta + b\gamma.$$

This trace is equal to the character of $g_1 g_2$ in the representation of the origin χ_2. In that manner we obtain the representation Γ_8 or Γ_9 according to the action of s. The representations Γ_1, Γ_2, Γ_3, and Γ_4 are obtained in the same way from the representations χ_1'' and χ_5''.

The table of characters thus obtained is identical with that given by Woodman (1966) without details of the construction.

REFERENCES

ALLEN, H. C., and CROSS, P. C. (1963). "Molecular Vib-Rotors." Wiley, New York.

ALTMANN, S. L. (1962). *In* "Quantum Theory" (D. R. Bates, ed.), Part II, Ch. 11. Academic Press, New York.

ALTMANN, S. L. (1967). *Proc. Roy. Soc. Ser. A* **298**, 184.

ALTMANN, S. L. (1971). *Mol. Phys.* **21**, 587.

BAK, B., CHRISTIANSEN, J. J., KUNSTMANN, K., NYGAARD, L., and RASTRUP-ANDERSEN, J. (1966). *J. Chem. Phys.* **45**, 883.

BARTON, D. H. R. (1970). *Angew. Chem.* **82**, 827.

BERRY, R. S. (1960). *J. Chem. Phys.* **32**, 933.

BIST, H. D., BRAND, J. C. D., and WILLIAMS, D. R. (1966). *J. Mol. Spectrosc.* **21**, 76.

BIST, H. D., BRAND, J. C. D., and WILLIAMS, D. R. (1967a). *J. Mol. Spectrosc.* **24**, 413.

BIST, H. D., BRAND, J. C. D., and WILLIAMS, D. R. (1967b). *J. Mol. Spectrosc.* **24**, 402.

BRAND, J. C. D., WILLIAMS, D. R., and COOK, T. J. (1966). *J. Mol. Spectrosc.* **20**, 359.

BRAND, J. C. D., CALIFANO, S., and WILLIAMS, D. R. (1968). *J. Mol. Spectrosc.* **26**, 398.

BRAND, J. C. D., JONES, V. T., FORREST, B. J., and PIRKLE, R. J. (1971). *J. Mol. Spectrosc.* **39**, 352.

BRIBES, J. L. (1971). Thesis, Univ. of Languedoc, Montpellier, France.

BRIBES, J. L., and GAUFRÈS, R. (1970). *J. Chim. Phys.* **67**, 1168.

BRIBES, J. L., and GAUFRÈS, R. (1971a). *Spectrochim. Acta, Part A* **27**, 2133.

BRIBES, J. L., and GAUFRÈS, R. (1971b). *J. Mol. Struct.* **9**, 423.

BUNKER, P. R. (1965a). *Mol. Phys.* **9**, 257.

BUNKER, P. R. (1965b). *J. Chem. Phys.* **42**, 2991.

BUNKER, P. R. (1967). *J. Chem. Phys.* **47**, 718.

BUNKER, P. R. (1968). *J. Chem. Phys.* **48**, 2832.

BUNKER, P. R., and HOUGEN, J. T. (1967). *Can. J. Phys.* **45**, 3867.

BUNKER, P. R., and LONGUET-HIGGINS, H. C. (1964). *Proc. Roy. Soc., Ser. A* **280**, 340.

BUNKER, P. R., and PAPOUŠEK, D. (1969). *J. Mol. Spectrosc.* **32**, 419.

CHOPLIN, F., and KAUFMANN, G. (1970). *Spectrochim. Acta, Part A* **26**, 2113.

CORIO, P. L. (1966). "Structure of High-Resolution NMR Spectra." Academic Press, New York.

DALTON, B. J. (1966). *Mol. Phys.* **11**, 265.

DEMAISON, J. (1972). Thesis, Univ. of Nancy I, Nancy, France.

DEMAISON, J., and RUDOLPH, H. D. (1971). *J. Mol. Spectrosc.* **40**, 445.

DREIZLER, V. H. (1961a). *Z. Naturforsch. A* **16**, 477.

DREIZLER, V. H. (1961b). *Z. Naturforsch. A* **16**, 1354.

DUNCAN, J. L. (1964). *Spectrochim. Acta* **20**, 1197.

DURIG, J. R., and LI, Y. S. (1972). *J. Mol. Struct.* **13**, 459.

FLEMING, J. W., and BANWELL, C. N. (1969). *J. Mol. Spectrosc.* **31**, 378.

FREEMAN, J. M., and HENSHALL, T. (1967). *J. Mol. Struct.* **1**, 31.

GILLES, J. M. F., and PHILIPPOT, J. (1972). *Int. J. Quantum. Chem.* **6**, 225.

GILLESPIE, P., HOFFMAN, P., KLUSACEK, H., MARQUARDING, D., PFOHL, S., RAMIREZ, F., TSOLIS, E. A., and UGI, I. (1971). *Angew. Chem. Int. Ed. Engl.* **10**, 687.

GUILLERMET, J., BRIBES, J. L., and GAUFRÈS, R. (1971). *J. Mol. Struct.* **9**, 97.

HAMADA, Y., HIRAKAWA, A. K., TAMAGAKE, K., and TSUBOI, M. (1970). *J. Mol. Spectrosc.* **35**, 420.

HARRIS, R. K., and THORLEY, M. (1972). *J. Mol. Spectrosc.* **42**, 407.

HIROTA, E. (1972). *J. Mol. Spectrosc.* **43**, 36.

HOUGEN, J. T. (1962). *J. Chem. Phys.* **37**, 1433.

HOUGEN, J. T. (1963). *J. Chem. Phys.* **39**, 358.

HOUGEN, J. T. (1964). *Can. J. Phys.* **42**, 1920.

HOUGEN, J. T. (1965). *Can. J. Phys.* **43**, 935.

HOUGEN, J. T. (1966). *Can. J. Phys.* **44**, 1169.

HOWARD-LOCK, H. E., and KING, G. W. (1970). *J. Mol. Spectrosc.* **35**, 393.

JOB, V. A., SETHURAMAN, V., and INNES, K. K. (1969). *J. Mol. Spectrosc.* **30**, 365.

KIRTMAN, B. (1964). *J. Chem. Phys.* **41**, 775.

KONDO, S., SAKURAI, Y., HIROTA, E., and MORONO, Y. (1970). *J. Mol. Spectrosc.* **34**, 231.

LAMB, D. (1966). Ph. D. Thesis, Oxford Univ., Oxford, England.

LIDE, D. R., and MANN, D. E. (1957). *J. Chem. Phys.* **27**, 874.

LONGUET-HIGGINS, H. C. (1963). *Mol. Phys.* **6**, 445.

MEYER, R., and GÜNTHARD, H. H. (1968). *J. Chem. Phys.* **49**, 1510.

MILLS, I. N. (1971). *Mol. Phys.* **20**, 127.

MUETTERTIES, E. L. (1970). *Accounts Chem. Res.* **3**, 266.

MYERS, R. I., and WILSON, E. B., JR. (1960). *J. Chem. Phys.* **33**, 186.

OLSON, W. B., and PAPOUŠEK, D. (1971). *J. Mol. Spectrosc.* **37**, 527.

PAPOUŠEK, D. (1968). *J. Mol. Spectrosc.* **28**, 161.

PAPOUŠEK, D., SARKA, K., SPIRKO, V., and JORDANOV, B. (1971). *Collect. Czech. Chem. Commun.* **36**, 890.

PETHRICK, R. A., and WYN-JONES, E. (1969). *Quart. Rev., Chem. Soc.* **23**, 301.

RUDOLPH, H. D., DREIZLER, H., JAESCHKE, A., and WENDLING, P. (1967). *Z. Naturforsch. A* **22**, 940.

SAUNDERS, J. E., LUCIER, J. J., and WILLIS, J. N., JR. (1968). *Spectrochim. Acta, Part A* **24**, 2023.

SERRE, J. (1967). *Int. J. Quantum Chem., Symp.* **1**, 713.

SERRE, J. (1968). *Int. J. Quantum Chem., Symp.* **2**, 107.

SERRE, J. (1969). *Atti Soc. Natur. Mat. Modena* 100, 69.

STONE, A. J. (1964). *J. Chem. Phys.* **41**, 1568.

SWALEN, J. D., and COSTAIN, C. C. (1959). *J. Chem. Phys.* **31**, 1562.

TANNENBAUM, E., MYERS, R. I., and GWINN, W. D. (1956). *J. Chem. Phys.* **25**, 42.

TUAZON, E. C., and FATELEY, W. G. (1971). *J. Chem. Phys.* **54**, 4450.

TURRELL, G. (1970). *J. Mol. Struct.* **5**, 245.

UGI, I., MARQUARDING, D., KUSACEK, H., and GILLESPIE, P. (1971). *Accounts Chem. Res.* **4**, 288.

WATSON, J. K. G. (1971). *Mol. Phys.* **21**, 577.

WIGNER, E. (1930). *Nachr. Ges. Wiss. Goettingen, Math.-Phys. Kl.* p. 133.

WILSON, E. B., DECIUS, J. C., and CROSS, P. C. (1955). "Molecular Vibrations." McGraw-Hill, New York.

WOODMAN, C. M. (1966). *Mol. Phys.* **11**, 109.

WOODMAN, C. M. (1970). *Mol. Phys.* **19**, 753.

Genealogical Electronic Spin Eigenfunctions and Antisymmetric Many-Electron Wavefunctions Generated Directly from Young Diagrams*

WILLIAM I. SALMON†

Ames Laboratory of the United States Atomic Energy Commission and
Department of Chemistry
Iowa State University
Ames, Iowa

* Based in part on a thesis submitted to Iowa State University in partial fulfillment of the requirements for the Ph.D. degree, 1971. Work was supported by the Ames Laboratory of the United States Atomic Energy Commission, Ames, Iowa, and the Chemical Physics Program, Washington State University, Pullman, Washington, under NSF Grant GU2631.

† *Present address:* Department of Physics, University of Utah, Salt Lake City, Utah.

I. Introduction

A. Indistinguishability of Electrons

Electrons are identical in the sense that no experiment can tell them apart. This implies that expectation values are independent of any electron numbering scheme. Suppose that $\Psi(1, 2, \ldots, N)$ is the exact wavefunction (a solution of the Schrödinger equation) for an N electron system, and that P is any of the $N!$ permutations of the electrons. Then for any observable operator $\mathscr{W}$,

$$\langle P\Psi(1, 2, \ldots, N) | \mathscr{W} P\Psi(1, 2, \ldots, N)\rangle$$
$$= \langle \Psi(1, 2, \ldots, N) | \mathscr{W}\Psi(1, 2, \ldots, N)\rangle.$$

Since permutations are unitary operators, it follows that

$$\langle \Psi(1, 2, \ldots, N) | P^{-1}\mathscr{W} P\Psi(1, 2, \ldots, N)\rangle$$
$$= \langle \Psi(1, 2, \ldots, N) | \mathscr{W}\Psi(1, 2, \ldots, N)\rangle$$

for any wavefunction Ψ. Thus it must be that $\mathscr{W} = P^{-1}\mathscr{W}P$: every observable operator is invariant under similarity transformations that permute its electron labels. In other words, every observable operator affects electrons symmetrically.

The behavior of the operators induces a behavior in their eigenfunctions. For if $\theta_j(N, W)$ is an N-electron eigenfunction of $\mathscr{W}$ such that

$\mathscr{W}\theta_j(N, W) = W\theta_j(N, W)$, then $P\theta_j(N, W)$ is an eigenfunction of $\mathscr{W}$ also, with the same eigenvalue:

$$\mathscr{W}[P\theta_j(N, W)] = P\mathscr{W}\theta_j(N, W) = W[P\theta_j(N, W)].$$

Since $[P\theta_j(N, W)]$ lies in the N-electron space for eigenvalue W, it can be written as a superposition of the $\mathscr{W}$-eigenfunctions $\{\theta_k(N, W) \mid k = 1, 2, \ldots, d(N, W)\}$ spanning that space:

$$P\theta_j(N, W) = \sum_{k=1}^{d(N,W)} \theta_k(N, W)[P]_{kj}^{NW}. \tag{1}$$

The numbers $[P]_{jk}^{NW}$ form a matrix $[P]^{NW}$ that depends in general on the choice of P and the values of N and W. In words, the eigenfunctions for a given eigenvalue of an observable operator span a representation of the symmetric group S_N of electronic permutations.

If this group contains *every* distinct symmetry transformation commuting with $\mathscr{W}$, then the degenerate functions $\{\theta_k\}$ span an *irreducible* representation of S_N (apart from accidental degeneracies), and each eigenvalue W is associated with a particular irreducible representation.

B. Exclusion Principle

Since permutations commute with the Hamiltonian, an implication of the argument above is that solutions of the N-electron Schrödinger equation for a given energy must span a representation of the symmetric group. Permutations of electrons do not in general comprise every distinct symmetry transformation commuting with a Hamiltonian, so there is no theoretical reason to suppose that such a representation will be irreducible.

Nevertheless, experiment demands that solutions of the Schrödinger equation for fermion systems must span the one-dimensional (thus irreducible) antisymmetric representation of the symmetric group. In other words, for every P in S_N,

$$P\Psi(1, 2, \ldots, N) = \varepsilon(P)\Psi(1, 2, \ldots, N),$$

where $\varepsilon(P)$ is $+1$ when P is even and -1 when P is odd. Here P is a transformation which simultaneously permutes the space *and* spin coordinates of the fermions.

This remarkable result, the Pauli exclusion principle for fermions, associates one particular irreducible permutational symmetry with all energy eigenvalues.

C. Spin Eigenfunctions

A second type of irreducible permutational symmetry is associated with spin eigenvalues. Spin angular momentum is very nearly a constant of the motion in many electronic systems of interest. Because of this, it is often desirable to construct wavefunctions to be eigenfunctions of the total spin operators, $\mathscr{S}^2$ and $\mathscr{S}_z$. This has the effect of "partially diagonalizing" observable matrices, simplifying the calculation of expectation values.

Let $\{\theta_j(N, S, M) \mid j = 1, 2, \ldots, d\}$ be a set of pure spin eigenfunctions with the properties

$$\mathscr{S}^2\theta_j(N, S, M) = \hbar^2 S(S+1)\theta_j(N, S, M)$$

and

$$\mathscr{S}_z\theta_j(N, S, M) = \hbar M\theta_j(N, S, M).$$

Since $\mathscr{S}^2$ and $\mathscr{S}_z$ are invariant under similarity transformations permuting the *spin coordinates* of electrons, we have from Eq. (1)

$$P^\sigma\theta_j(N, S, M) = \sum_{k=1}^{d} \theta_k(N, S, M)[P]_{kj},$$

where P^σ denotes an electronic spin permutation. Thus a set of spin eigenfunctions belonging to the given values of N, S, and M span a representation of the symmetric group $S_N{}^\sigma = \{P^\sigma\}$.

When the properties of such a *spin representation* are investigated (Weyl, 1931; Wigner, 1931; Kotani *et al.*, 1955), it is found that: (1) the dimension d of the representation and the representation matrix $[P]$ itself depend on N and S, but not on M; (2) the representation is irreducible.

The latter statement is true because $S_N{}^\sigma$ contains every distinct symmetry transformation commuting with the total spin operator $\mathscr{S}^2$. As Dirac (1929) has shown, $\mathscr{S}^2$ can be written in terms of class sums in $S_N{}^\sigma$.

The close relation between spin and the symmetric group leads to an alternative definition of spin eigenfunctions: they are basis functions for certain "irreps" (irreducible representations) of $S_N{}^\sigma$.

D. Spin-Adapted Antisymmetrized Wavefunctions

We have mentioned the usefulness of approximate electronic wavefunctions that are constructed not only to satisfy the Pauli principle, but also to be eigenfunctions of $\mathscr{S}^2$ and $\mathscr{S}_z$. Such "spin-adapted" antisymmetrized wavefunctions can always be expressed as superpositions of functions of the form

$$\Psi_k(N, S, M) = \mathscr{A}[\phi(N)\theta_k(N, S, M)],$$

where $\mathscr{A}$ is the N-electron antisymmetrizer, $\phi(N)$ is a space function of some sort, and $\theta_k(N, S, M)$ is a pure spin eigenfunction.

Several methods have been popular for constructing the spin functions. The earliest was the spin-pairing scheme evolved by Heitler and London (1927), Slater (1931), Pauling (1933), Rumer (1932), and McLachlan (1960) for use in the valence–bond approximation. This scheme was later generalized and recast in elegant form by Matsen and co-workers (Matsen, 1964; Matsen *et al.*, 1966). One obtains linearly independent, but not orthogonal, spin functions.

More recently, it has become popular to use Löwdin's (1955, 1960, 1964) projection operator technique to extract spin eigenfunctions from pure spin primitives, such as spin products. This approach has a directness that lends itself to computer calculations; it can be extended to other angular momenta as well. But again, the spin eigenfunctions are not orthogonal.

In many applications it is most convenient to construct wavefunctions from orthogonal spin functions generated by the so-called "genealogical," or spin-coupling, techniques. Other papers in this series deal with the considerable advantages of such functions in unrestricted configuration–interaction calculations on electronic systems (Ruedenberg, 1971; Salmon and Ruedenberg, 1972a) and with their construction by methods not involving group theory (Salmon *et al.*, 1972; Salmon and Ruedenberg, 1972b). The present paper uses the powerful techniques of symmetric group theory to obtain simple schemes for generating genealogical spin functions and antisymmetrized spin-adapted wavefunctions.

E. Outline

As we have seen, antisymmetrized spin-adapted wavefunctions are characterized by two kinds of permutational symmetry: (1) they are antisymmetric with respect to simultaneous permutations of the space and spin coordinates of the electrons (the exact meaning of "antisymmetrized"); (2) they transform according to certain irreducible representations of the symmetric group permuting only the spin coordinates of the electrons (the group-theoretical significance of spin-adaptation).

As these properties suggest, it is possible to generate spin-adapted antisymmetrized wavefunctions from space and spin primitive functions (such as orbital and spin products) by means of group-algebraic operators. Such an approach is especially convenient *when the operators can be written down directly from Young diagrams and standard tableaux.*

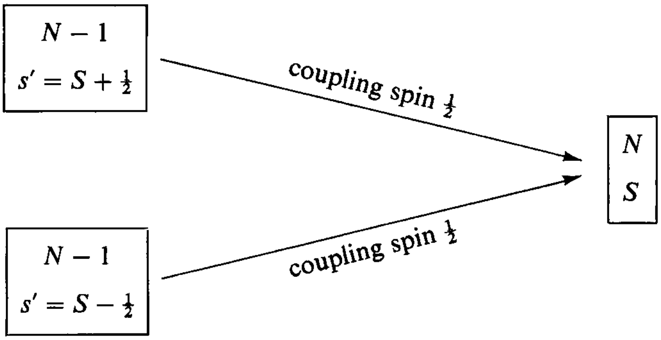
The present paper introduces a method by which the information embodied in Young diagrams and tableaux can be used to *write down at sight* the operators required to generate (1) genealogical spin functions of the two most common types, and (2) spin-adapted antisymmetrized wavefunctions defined in terms of such spin functions. While Young operators have been used in the past to generate the "spin-paired" functions of valence–bond theory (Matsen, 1964; Matsen *et al.*, 1966), it has heretofore been impossible to generate genealogical or other orthogonal spin functions in this way without brute-force orthogonalization. The method developed in the following pages is more elegant and seems well suited to high-speed computing.

The two most common types of genealogical spin functions are discussed in Section II. In Section III, general methods are described for generating these functions by means of Wigner operators, defined either in terms of representation matrices or as operator bases in the symmetric group algebra. Section IV introduces the novel material: a heuristic argument is used to find schemes by which genealogical Wigner operators can be written down at sight from Young tableaux. The validity of the schemes is proved in Section V. In the concluding part, the nature of the space–spin duality is examined from the points of view of classical group theory and group algebra theory. It is shown that the operators introduced in this paper lead to a simple form for antisymmetrized spin-adapted wavefunctions.

II. Genealogical Spin Functions

A. Yamanouchi–Kotani Functions

The best-known genealogical construction for electronic spin states is that invented by Yamanouchi (1936, 1937, 1938) and elaborated by Kotani *et al.* (1955). Here spins are coupled one at a time: one starts with the spin of a single electron, couples it to the spin of another, and proceeds by coupling the spin of the Nth electron to the resultant spin of the first $(N - 1)$. At each stage, there are two ways in which one can obtain spin S for N electrons. Pictorially,



This sort of spin-coupling picture is called a *branching diagram*, and the two *routes* shown correspond to the two equations

$$\theta_j(N,S,M) = -\left(\frac{S-M+1}{2S+2}\right)^{1/2} \theta_j(N-1, S+\tfrac{1}{2}, M-\tfrac{1}{2})\cdot\alpha(N) + \left(\frac{S+M+1}{2S+2}\right)^{1/2} \theta_j(N-1, S+\tfrac{1}{2}, M+\tfrac{1}{2})\cdot\beta(N) \quad \text{(2a)}$$

and

$$\theta_j(N,S,M) = \left(\frac{S+M}{2S}\right)^{1/2} \theta_j(N-1, S-\tfrac{1}{2}, M-\tfrac{1}{2})\cdot\alpha(N) + \left(\frac{S-M}{2S}\right)^{1/2} \theta_j(N-1, S-\tfrac{1}{2}, M+\tfrac{1}{2})\cdot\beta(N). \quad \text{(2b)}$$

The coefficients appearing here are examples of Clebsch–Gordan or Wigner coefficients (Wigner, 1931) and guarantee that the $\theta_j(N, S, M)$ form an orthonormal basis for an irreducible representation of S_N.

In applying these equations recursively for given N, S, and M, one makes a spin-coupling choice at each stage—a choice between Eqs. (2a) and (2b). In the end, there are a number of ways in which N one-electron spins can be coupled so that the resultant spin is S. Each of these "spin-coupling schemes" involves constructing a spin function *from a chain of predecessors*, and is labeled by a value of the subscript j in Eq. (2). The various schemes can be represented pictorially as routes on an N-electron branching diagram like the one in Fig. 1, where at each intersection is shown the number of spin functions resulting for the corresponding values of N and S. This number (the dimension of the irrep of S_N labeled by N and S) is independent of M, with the value (Kotani *et al.*, 1955)

$$d(N,S) = \frac{(2S+1)(N!)}{(N/2+S+1)!\,(N/2-S)!} = \frac{(2S+1)}{(N+1)}\binom{N+1}{N/2-S}.$$

Thus, for example, there are three spin eigenfunctions for $N = 4$, $S = 1$, *for each value of M*.

The Yamanouchi-Kotani spin functions (or YK functions, as we shall hereafter refer to them) are a basis for a very special orthogonal representation of S_N. First of all, the representation of S_N itself is irreducible since $\theta_j(N, S, M)$ is, for each value of j, an eigenfunction of the total spin operator for N electrons. But it is clear from Eq. (2) that $\theta_j(N, S, M)$ is also an eigenfunction of the total spin operators for $(N - 1)$ electrons, for $(N - 2)$ electrons, ..., and for one electron. Therefore, the recursive nature of these equations has the result that the representations of the

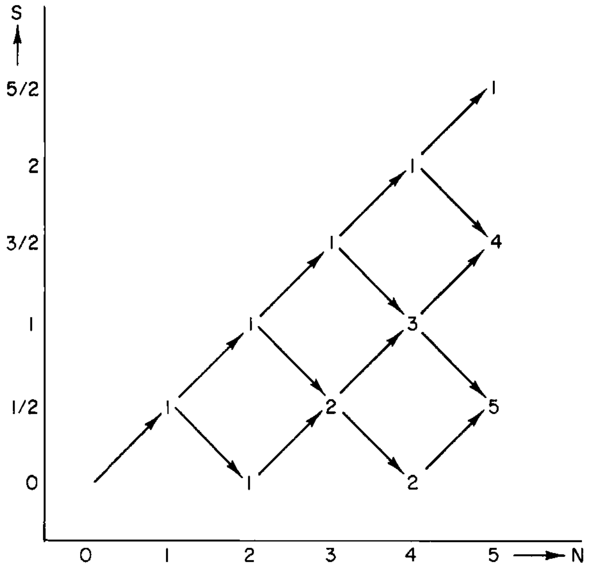


Fig. 1. Yamanouchi–Kotani branching diagram.

subgroups $S_{N-1}, S_{N-2}, \ldots, S_1$ are *all* fully reduced. The YK representation is said to be *adapted to the sequence* $S_N \supset S_{N-1} \supset S_{N-2} \supset \cdots \supset S_1$ of nested symmetric groups (Klein *et al.*, 1970). We shall return to this point later.

B. Serber Functions

In previous papers of this series (Ruedenberg, 1971; Salmon and Ruedenberg, 1972a,b; Salmon *et al.*, 1972), we discussed the usefulness of orthogonal spin eigenfunctions that are simultaneously eigenfunctions of all the " geminal spin operators" $\mathscr{S}_\mu{}^2$, where μ labels the "geminal pair' of electrons $2\mu - 1$ and 2μ. Such functions were first obtained by Serber (1934a,b) using a genealogical procedure in which spins were coupled *two* at a time.

Suppose that $N = 2n$ is even. Then, defining geminal spin functions $w_\mu(s_\mu, m_\mu)$ for the μth geminal pair of electrons,

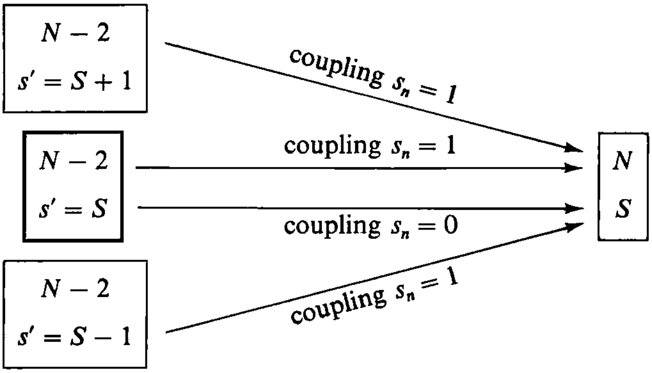
$$\begin{aligned} w_\mu(1, 1) &= \alpha(2\mu - 1)\alpha(2\mu), \\ w_\mu(1, 0) &= [\alpha(2\mu - 1)\beta(2\mu) + \beta(2\mu - 1)\alpha(2\mu)]/\sqrt{2}, \\ w_\mu(1, -1) &= \beta(2\mu - 1)\beta(2\mu), \\ w_\mu(0, 0) &= [\alpha(2\mu - 1)\beta(2\mu) - \beta(2\mu - 1)\alpha(2\mu)]/\sqrt{2}, \end{aligned} \tag{3}$$

one constructs $2n$-electron functions stepwise, using equations analogous to Eq. (2):

$$\theta_j'(N, S, M) = \sum_{m_n} W_j(s', s_n, S; M - m_n, m_n, M) \cdot \theta_j'(N - 2, s', M - m_n) w_n(s_n, m_n). \tag{4}$$

Here $\theta_j'(N - 2, s', M - m_n)$ is an $(N - 2)$-electron spin function for spin s' and z-component $M - m_n$. Since s_n (the spin of the nth geminal pair) can be 0 or 1, s' can be $S + 1$, S, or $S - 1$. The numbers $W_j(s', s_n, S; M - m_n, m_n, M)$ are Wigner coupling coefficients.

There are four equations like Eq. (4), corresponding to the four spin-coupling ("branching") routes shown in the following diagram:



The different subscripts j occurring in Eq. (4) correspond to different routes on a Serber branching diagram like that in Fig. 2. As in the previous case, the values of $d(N, S)$ are shown at each intersection.

It follows from Eq. (4) that Serber spin functions are a basis for a representation of S_N that is adapted to the sequence $S_N \supset S_{N-2} \supset S_{N-4} \supset \cdots \supset S_2$ of nested symmetric groups. It also follows from this equation that the representation of every geminal two-electron subgroup is fully reduced. These facts prove useful in certain applications (Ruedenberg, 1971; Salmon and Ruedenberg, 1972a; Miller and Ruedenberg, 1968).

"Serber-type" functions for *odd* N can be made by coupling the spin of the Nth electron to Serber functions for $N' = N - 1$. The resulting functions will then have Serber-type behavior up to electron N'.

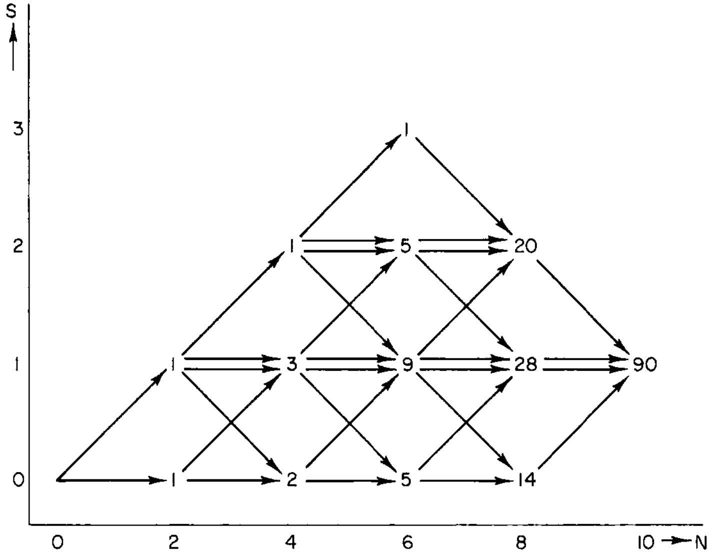


Fig. 2. Serber branching diagram for states leading to $N = 10$, $S = 1$.

C. Comparison of YK and Serber Functions

The differences between YK and Serber spin functions may not be obvious from the branching diagrams, Figs. 1 and 2. The easiest way to reveal the differences is to examine the functions resulting from both genealogical schemes when, say, $N = 4$, $S = 1$, $M = 0$. We use the notation introduced previously and show with each function its branching route.

The YK functions turn out to be

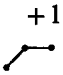
 : $\theta_1(410) = (\alpha\beta - \beta\alpha)(\alpha\beta + \beta\alpha)/2$,

 : $\theta_2(410) = [2\alpha\alpha\beta\beta - 2\beta\beta\alpha\alpha - (\alpha\beta + \beta\alpha)(\alpha\beta - \beta\alpha)]/2\sqrt{3}$,

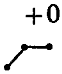
 : $\theta_3(410) = [\alpha\alpha\beta\beta - \beta\beta\alpha\alpha + (\alpha\beta + \beta\alpha)(\alpha\beta - \beta\alpha)]/\sqrt{6}$.

On the other hand, the Serber functions are

 : $\theta'_1(410) = (\alpha\beta - \beta\alpha)(\alpha\beta + \beta\alpha)/2$,



: $\theta'_2(410) = (\alpha\alpha\beta\beta - \beta\beta\alpha\alpha)/\sqrt{2}$,



: $\theta'_3(410) = (\alpha\beta + \beta\alpha)(\alpha\beta - \beta\alpha)/2$.

The Serber functions are symmetric or antisymmetric in each geminal pair: they are simultaneous eigenfunctions of $\mathscr{S}^2$, $\mathscr{S}_z$, ${\mathscr{S}_1}^2$, and ${\mathscr{S}_2}^2$. The YK functions are less simple. The first one happens to be the same as the Serber function θ_1' because its branching diagram unambiguously fixes the spin of the first geminal pair to be $s_1 = 0$. Since the total spin is one, the spin of the second pair must be $s_2 = 1$. In the other two YK functions, the spin of the first geminal pair is unambiguously $s_1 = 1$, but the second pair has no definite spin. In other words, the functions θ_2 and θ_3 are simultaneous eigenfunctions of $\mathscr{S}^2$, $\mathscr{S}_z$, and ${\mathscr{S}_1}^2$, but not of ${\mathscr{S}_2}^2$. This is because either $s_2 = 1$ or $s_2 = 0$ can couple with $s_1 = 1$ to give $S = 1$. Rather than containing a pure contribution from $s_2 = 1$ or $s_2 = 0$, the YK functions θ_2 and θ_3 contain mixtures of both.

D. Practicality of Spin-Coupling Techniques

The genealogical construction of spin functions is inconvenient because it is recursive. In order to make an N-electron spin function, one must first generate every predecessor in the genealogical scheme. It can be seen from the branching diagrams that the complexity of the problem increases rapidly with N.

In order to make the three YK functions for $N = 4$, $S = 1$, $M = 0$, one must generate the fifteen functions given in Table I. The calculations are so simple that this is no problem. But there are 90 spin functions for $N = 10$, $S = 1$, $M = 0$. In order to get them, one must generate 660 functions altogether, some containing as many as 252 product functions.

Computer programs have been made for the genealogical construction of YK functions (Mattheiss, 1958) and Serber permutation representation matrices (Mattheiss, 1959). However, such programs require an inconvenient amount of storage for any but small N. As an alternative approach, we discuss here some group-theoretical techniques for obtaining genealogical spin functions.

TABLE I

FIFTEEN YK FUNCTIONS GENERATED FOR $N = 4$, $S = 1$, $M = 0$

N	S	Functions for each M	M values required	Total functions
1	$\frac{1}{2}$	1	$+\frac{1}{2}, -\frac{1}{2}$	2
2	0	1	0	1
2	1	1	$1, 0, -1$	3
3	$\frac{1}{2}$	2	$+\frac{1}{2}, -\frac{1}{2}$	4
3	$\frac{3}{2}$	1	$+\frac{1}{2}, -\frac{1}{2}$	2
4	1	3	0	3

III. Spin Functions by Group-Theoretical Techniques

A. Wigner Operator Bases for Orthogonal Irreps of S_N

Basis functions for any given irrep of S_N can be generated from a primitive function by means of the appropriate *Wigner operators*. Let $[P]^\alpha$ be the matrix representing the permutation P in the irrep α of dimension d^α. Then the operators

$$e_{rs}^{\alpha} = \frac{d^{\alpha}}{N!} \sum_{P \in S_N} [P^{-1}]_{sr}^{\alpha} P, \qquad r = 1, 2, \ldots, d^{\alpha}, \tag{5}$$

where s is fixed, are linearly independent and span a carrier space for the irrep (Wigner, 1931).

It is easily verified that

$$Pe_{rs}^{\alpha} = \sum_{t} [P]_{tr}^{\alpha} e_{ts}^{\alpha},$$

showing that different choices of s in Eq. (5) produce different bases for the same irrep. We say that the Wigner operators $\{e_{rs}^{\alpha} \mid \alpha, s \text{ fixed}\}$ form an *operator basis* for the irrep α, as opposed to the *function basis* they can be used to generate.

It is not practical to generate basis functions by means of definition (5), except when N is relatively small. In the first place, it is a formidable task to evaluate all the irrep matrix elements required. Even if this is done, Eq. (5) involves summing over $N!$ permutations, and so is impractical when $N \gtrsim 8$. *Our aim here is to find simple rules whereby operator bases for YK and Serber representations can be written down directly from Young diagrams.* We begin by investigating the connection between Wigner and Young operators.

Wigner operators in S_N are linear combinations of permutations of the form

$$X = \sum_{P} \xi(P)P, \tag{6}$$

where $\xi(P)$ is some number that depends on the permutation P. In other words, Wigner operators are particular *elements of the symmetric group algebra*. It is from this point of view that we shall consider them.

Given two well-behaved functions ψ and ϕ, we observe that operators X in the group algebra have the following integral property:

$$\begin{aligned}\langle X\psi \mid \phi \rangle &= \sum_{P} \xi^*(P)\langle P\psi \mid \phi\rangle = \sum_{P} \xi^*(P)\langle \psi \mid P^{-1}\phi\rangle \\ &= \langle \psi \mid \sum_{P} \xi^*(P)P^{-1}\phi\rangle.\end{aligned}$$

The adjoint of X is therefore defined to be

$$X^{\dagger} = \sum_{P} \xi^{*}(P)P^{-1}.$$

The particular Wigner operators we seek to construct have real coefficients which are elements of orthogonal matrices. It follows that these operators have the adjoint property

$$\begin{aligned} e_{rs}^{\alpha\dagger} &= \kappa^{\alpha} \sum_{P} [P^{-1}]_{sr}^{\alpha} P^{-1} = \kappa^{\alpha} \sum_{P} [P]_{rs}^{\alpha} P^{-1} \\ &= \kappa^{\alpha} \sum_{P} [P^{-1}]_{rs}^{\alpha} P = e_{sr}^{\alpha}. \end{aligned} \tag{7}$$

Conversely, it can be shown that this property guarantees an orthogonal representation.

By application of the orthogonality theorem for irreducible representations, it is easily seen that the Wigner operators defined by Eq. (5) possess the multiplicative property

$$e_{rs}^{\alpha} e_{tu}^{\beta} = \delta^{\alpha\beta} \delta_{st} e_{ru}^{\alpha}. \tag{8}$$

Thus these operators, of which there are $\sum_{\alpha} (d^{\alpha})^{2} = N!$ altogether, are all linearly independent and span the space of all linear operators of form (6). In other words, *they are a basis for the entire symmetric group algebra,* $A(S_N)$. Because of this and the fact that Eq. (8) resembles the multiplicative property of "elementary matrices," the e_{rs}^{α} are often referred to as a *matric basis* (Matsen, 1964; Matsen *et al.*, 1966; Littlewood, 1950). We deal here only with matric bases having the adjoint property (7): ones that are associated with orthogonal representations. These we shall call *orthogonal matric bases.*

The diagonal elements e_{rr}^{α} of an orthogonal matric basis are especially interesting. From Eq. (8), these operators have the property

$$e_{rr}^{\alpha} e_{ss}^{\beta} = \delta^{\alpha\beta} \delta_{rs} e_{rr}^{\alpha}: \tag{9}$$

they are idempotent and annihilate each other from the left and right. From this it is easy to show that they form a resolution of the group identity I:

$$\sum_{\alpha} \sum_{r} e_{rr}^{\alpha} = I. \tag{10}$$

In addition, it follows from Eq. (7) that the diagonal elements of an orthogonal matric basis are *Hermitian*:

$$e_{rr}^{\alpha\dagger} = e_{rr}^{\alpha}. \tag{11}$$

Taken together, properties (9)–(11) show that the e_{rr}^{α} are *projection operators* for the irreducible carrier spaces into which the group algebra decomposes. Since the group algebra is a carrier space for the regular representation, it is seen that the set $\{e_{rr}^{\alpha}\}$ contains projection operators for *all* of the distinct (i.e., nonequivalent) irreducible representations of the symmetric group.

B. Minimal Left Ideals and Primitive Idempotents

In order to proceed with the discussion, it is necessary to use the special terminology of group algebra theory. We give here a brief summary of the material required later. Detailed and complete treatments are available elsewhere (Weyl, 1931; Boerner, 1963; Löwdin, 1967).

Given an element e in the group algebra $A(S_N)$, the subset

$$L = \{Xe;\ X \in A(S_N)\}$$

is said to be the *left ideal* generated by e. If e is idempotent, it is said to be a *generating unit* for L. It can be shown that every left ideal possesses at least one generating unit.

By definition, a left ideal is a subspace of the group algebra that is invariant under left multiplication by elements of that algebra. But the group algebra is a carrier space for the regular representation. Therefore, a left ideal is a carrier space for one of the representations into which the regular representation reduces. Just as the regular representation ultimately reduces to a direct sum of all the distinct irreducible representations of S_N, the group algebra ultimately *decomposes* (reduces) to a direct sum of certain *minimal* (irreducible) left ideals:

$$A(S_N) = L_1 \oplus L_2 \oplus \cdots \oplus L_k. \tag{12}$$

Each L_i is an operator carrier space for an irrep of S_N. The irrep α, of dimension d^α, is represented in the direct sum by d^α equivalent minimal left ideals. To emphasize this point, we rewrite Eq. (12) in the form

$$A(S_N) = \sum_\alpha (L_1{}^\alpha \oplus L_2{}^\alpha \oplus \cdots \oplus L_{d^\alpha}^\alpha), \tag{13}$$

the summation being direct. This decomposition is uniquely defined except for order and equivalence.

The generating units for *minimal* left ideals are called *primitive idempotents*. These are, then, operators that generate carrier spaces for irreducible representations. If the operators are self-adjoint, as in Eq. (11), their minimal left ideals are carrier spaces for *orthogonal* irreducible representations. A set of primitive idempotents annihilating each other according to Eq. (9) and resolving the identity, as in Eq. (10), generates a set of minimal left ideals forming a decomposition of the group algebra. *Thus, in group-algebraic parlance, the diagonal Wigner operators* $\{e_{rr}^\alpha\}$ *are primitive idempotents generating a set of minimal left ideals into which the group algebra decomposes.*

Group algebra theory provides a means by which an operator can be identified as a primitive idempotent. According to a well-known theorem, an element e of the group algebra $A(S_N)$ is a primitive idempotent if and only if

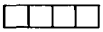
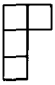
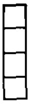
$$eXe = \lambda(X)e, \tag{14}$$

for *every* element X in $A(S_N)$, where $\lambda(X)$ is a number that depends on X. (This number may be zero for some choices of X, but cannot be zero for every X unless e is trivially the null.) As we shall see, group algebra theory also provides a diagrammatic procedure by which one can generate operators satisfying not only Eq. (14), but also Eqs. (9)–(11).

C. Young Diagrams, Tableaux, and Operators

Minimal left ideals in the symmetric group algebra can be generated by a method developed by Young (1901, 1902, 1928, 1930, 1932). Accounts of this technique have been given by Rutherford (1948), Weyl (1931), and Boerner (1963). The following summary is intended merely as a glossary of terminology and notation to be used later.

There are as many classes of the group S_N as there are partitions of the number N. On the other hand, there are as many classes of a group as there are distinct irreps. Consequently, the partitions of N provide a means for labeling the irreps of S_N. These partitions are often represented by "pictures" called *Young diagrams*. For example, the five irreps for S_4 can be labeled as follows:

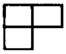
Partition	Young diagram
$\{4\}$	
$\{3, 1\}$	
$\{2, 2\}$	
$\{2, 1, 1\} = \{2, 1^2\}$	
$\{1, 1, 1, 1\} = \{1^4\}$	

Thus, a diagram with κ rows corresponds to the partition

$$\{\rho_1, \rho_2, \ldots, \rho_\kappa\} \qquad (\rho_1 \geq \rho_2 \geq \cdots \geq \rho_\kappa),$$

where ρ_i is the number of boxes on row i.

Each diagram is used to construct *Young tableaux*. A tableau is a particular arrangement of the numbers 1, 2, ..., N in the boxes of a diagram. If the numbers increase to the right along rows and downward along columns, the tableau is said to be *standard*. For example, the diagram



for $N = 3$ gives rise to the six tableaux

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 3 & 2 \\ \hline 1 \\ \cline{1-1}\end{array},$$

of which only the first two are standard.

In the following we deal only with standard tableaux. It can be shown that there are d^D of these, where d^D is the dimension of the irrep labeled by diagram D. We label the standard tableaux for diagram D $\{T_r^D;\ r = 1, 2, \ldots, d^D\}$. Each standard tableau labels one of the equivalent irreducible subspaces for irrep "D" occurring in the regular representation. The problem is to generate bases in these subspaces.

The *row group*, $\mathscr{R}_r^D$, for the standard tableau T_r^D consists of all permutations which interchange only numbers on the same row of T_r^D. This group is the direct product of the row groups for each row, taken separately. A *column group*, $\mathscr{C}_r^D$, is defined analogously with reference to the columns.

Operators are formed for each standard tableau. The *row operator* for T_r^D is an unnormalized symmetrizer over the row group:

$$R_r^D = \sum_{r \in \mathscr{R}_r^D} r.$$

The column operator for T_r^D is an unnormalized *anti*symmetrizer over $\mathscr{C}_r^D$:

$$C_r^D = \sum_{c \in \mathscr{C}_r^D} \varepsilon(c)c, \qquad \text{where } \varepsilon(c) = +1 \text{ when } c \text{ is even}$$
$$= -1 \text{ when } c \text{ is odd}.$$

Finally, the "tableau operator" is defined to be the column operator followed by the row operator:

$$E_r^D = R_r^D C_r^D. \tag{15}$$

(Some authors define $E_r^D = C_r^D R_r^D$.)

The "tableau operator" E_r^D is the foundation for the material that follows. It can be shown to satisfy Eq. (14):

$$E_r^D X E_r^D = (\text{number}) \cdot E_r^D,$$

for every element X in $A(S_N)$. In particular, when X is the identity,

$$E_r^D E_r^D = (\text{number}) \cdot E_r^D.$$

Thus E_r^D is, within a numerical factor, a primitive idempotent, and the set $\{XE_r^D;\ X \in A(S_N)\}$ is a minimal left ideal. The operator E_r^D is called a *Young idempotent.*

Application of a Young idempotent to a primitive function f produces a function symmetry-adapted to an irrep of S_N. This means that

(1) For any primitive function f the function $(E_r^D f)$, if it is not trivially zero, belongs to a carrier space for the irreducible representation of S_N labeled by the diagram D.

(2) Given any linear combination X of permutations, the function $(XE_r^D f)$ belongs to the same irreducible carrier space as $(E_r^D f)$.

(3) The entire irreducible carrier space can be generated by sweeping X over the group algebra.

Suppose that the irrep associated with Young diagram D has dimension d^D, and that p_{sr}^D denotes the permutation that transforms the standard tableau T_r^D into T_s^D:

$$T_s^D = p_{sr}^D T_r^D. \tag{16}$$

Any set of d^D linearly independent operators of the form XE_r^D will be an operator basis for the irrep. Young showed that the operators

$$E_{sr}^D = p_{sr}^D E_r^D \qquad (r \text{ arbitrary};\quad s = 1, 2, \ldots, d^D),$$

form such a basis. Basis functions for the irrep associated with diagram D can then be generated from a suitable primitive function by applying the *Young operators* $\{E_{sr}^D;\ s = 1, 2, \ldots, d^D\}$. Different choices of r give different bases for the same representation.

It should be noted that the relations (16) between tableaux imply the following relations between operators defined in terms of the tableaux:

$$R_s^D = p_{sr}^D R_r^D p_{rs}^D, \qquad C_s^D = p_{sr}^D C_r^D p_{rs}^D, \qquad E_s^D = p_{sr}^D E_r^D p_{rs}^D,$$

where, clearly, $p_{rs}^D = (p_{sr}^D)^{-1}$ and $p_{rr}^D = I =$ identity. Consequently, the Young operators are most generally defined by

$$E_{sr}^D = p_{sr}^D R_r^D C_r^D = R_s^D p_{sr}^D C_r^D = R_s^D C_s^D p_{sr}^D. \tag{17}$$

We shall make much use of these relations shortly.

D. Electronic Spin Eigenfunctions by Means of Young Operators

Any operator antisymmetrizing more than two numbers will annihilate an electronic spin function, which contains only two distinct spinors—α and β. It follows that the column operator C_r^D, and thus the Young operator $E_{sr}^D = p_{sr}^D R_r^D C_r^D$, will annihilate a spin primitive if T_r^D contains more than two rows. Clearly, electronic spin eigenfunctions span irreps of S_N associated with Young diagrams of only one or two rows. We shall refer to these as *spin diagrams*.

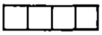
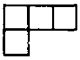
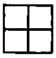
The correlation between Young spin diagrams and spin states (N, S) is determined by checking the dimensionality of the irrep—the number of standard tableaux. Comparing this number with the spin degeneracy $d(N, S)$ [Figs. 1 and 2], it is found that the spin diagram $D = \{\rho_1, \rho_2\}$ is associated with the spin state having

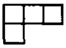
$$N = \rho_1 + \rho_2, \qquad S = (\rho_1 - \rho_2)/2.$$

For example, the correlation for four electrons is given in Table II.

TABLE II

SPIN DIAGRAMS AND SPIN STATES FOR FOUR ELECTRONS

Young spin diagram	Number of standard tableaux = $d(N, S)$	Spin quantum number
	1	$S = 2$
	3	$S = 1$
	2	$S = 0$

As an example of spin functions generated by means of Young operators, we may consider the case of four electrons with $S = 1$. The Young diagram is $D =$  , for which there are three standard tableaux:

$$T_1^D = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1}\end{array}, \qquad T_2^D = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array}, \qquad T_3^D = \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 \\ \cline{1-1}\end{array}.$$

Denoting by $\mathscr{S}_{ij\ldots k}$ and $\mathscr{A}_{ij\ldots k}$ the unnormalized symmetrizer and antisymmetrizer on the numbers $\{i, j, \ldots, k\}$, an operator basis for the irrep is given by the Young operators $\{E_{i1}^D = p_{i1}^D R_1{}^D C_1{}^D;\ i = 1, 2, 3\}$, where $p_{11}^D = I$, $p_{21}^D = (3, 4)$, $p_{31}^D = (2, 3)(3, 4)$, and $R_1{}^D C_1{}^D = E_{11}^D = \mathscr{S}_{123}\mathscr{A}_{14}$. Using the spin primitive function $f = \alpha\beta\alpha\beta$, we obtain the spin eigenfunctions (except for a numerical factor)

$$\begin{aligned}
E_{11}^D f &= \alpha\alpha\beta\beta - \beta\beta\alpha\alpha + (\alpha\beta + \beta\alpha)(\alpha\beta - \beta\alpha),\\
E_{21}^D f &= (3, 4)E_1{}^D f = \alpha\alpha\beta\beta - \beta\beta\alpha\alpha - (\alpha\beta + \beta\alpha)(\alpha\beta - \beta\alpha),\\
E_{31}^D f &= (2, 3)(3, 4)E_1{}^D f = -\alpha\alpha\beta\beta + \beta\beta\alpha\alpha + (\alpha\beta - \beta\alpha)(\alpha\beta + \beta\alpha).
\end{aligned}$$

These functions are linearly independent but not orthogonal. In addition, they do not display the subgroup adaptation typical of YK or Serber spin functions. They are not suited to the sort of applications discussed in other papers of this series. The inadequacy, of course, is in the operators.[1]

E. Connection between Wigner and Young Operators

In the remaining sections of this paper, Young idempotents will be used to construct matric bases—Wigner operators—for the orthogonal YK and Serber representations of S_N. It is appropriate at this point to make a reconnaissance.

As mentioned previously, the diagonal elements of matric bases multiply according to

$$e_{rr}^D e_{ss}^{D'} = \delta^{DD'}\delta_{rs} e_{rr}^D.$$

That is, these elements are idempotent and mutually annihilating. According to group algebra theory, this property is required of generating units for the minimal left ideals into which the group algebra decomposes. Such primitive idempotents are also required to have the property (14).

Young idempotents are primitive—they satisfy Eq. (14). Also, two Young idempotents from different diagrams annihilate each other from the left and right. But it can be shown that two Young idempotents from the *same* diagram may not do this. Thus Young idempotents generate minimal left ideals, but do not quite multiply like the diagonal elements of a matric basis (McIntosh, 1960).

The situation is as follows. In the decomposition of $A(S_N)$ there occur d^D equivalent irreducible carrier spaces—minimal left ideals—for the irrep labeled by diagram D. These carrier spaces are generated by the matric basis idempotents $e_{11}^D, e_{22}^D, \ldots, e_{d^D d^D}^D$. The Young idempotents $E_1{}^D, E_2{}^D, \ldots, E_{d^D}^D$ also generate carrier spaces for this representation.

---

[1] Young operators can be used to generate the "spin-paired" functions in the valence bond approximation (see Matsen, 1964; Matsen *et al.*, 1966).

Thus there must be equivalence transformations relating the Young idempotents and the matric basis idempotents.

Furthermore, the Young operators E_{sr}^D do not possess the adjoint property (7) associated with an orthogonal matric basis. Instead,

$$E_{sr}^{D\dagger} = C_r^{D\dagger} R_r^{D\dagger} p_{sr}^{D\dagger} = C_r{}^D R_r{}^D p_{rs}^D = p_{rs}^D C_s{}^D R_s{}^D \neq E_{rs}^D .$$

Finally, there is nothing about the construction of these operators that would associate them with a genealogical spin-coupling scheme.

We shall construct from Young idempotents new sets of operators that (1) multiply like matric bases, (2) have the adjoint property of Eq. (7), and (3) are related to genealogical spin-coupling schemes. This will be accomplished by multiplying Young idempotents from the left and right by certain operators.

IV. Generating Genealogical Spin Functions from Young Diagrams

A. Tableau Chains

It is well known that standard tableaux can be derived from a genealogical scheme similar to that involved in spin coupling (Klein *et al.*, 1970; Jahn and van Wieringen, 1951; Pauncz, 1967; Coleman, 1968). Since Young spin diagrams $\{\rho_1, \rho_2\}$ label spin representations of S_N through the relations $\rho_1 = (N/2) + S$, $\rho_2 = (N/2) - S$, the YK branching diagram can be given in the form shown in Fig. 3. In other words, the Young diagrams

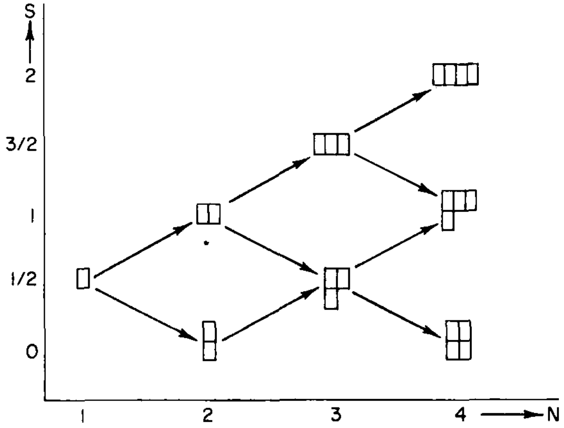


Fig. 3. YK branching diagram for Young diagrams.

can be considered the result of a " box-coupling " procedure: one starts with $\square$ and adds boxes one by one, subject to the condition that $\rho_1 \geq \rho_2$.

Figure 3 is a kind of shorthand for the genealogical construction of standard tableaux. If we start with the tableau $\boxed{1}$ and add, one by one, the numbers 2, 3, ..., N in such a way that the resulting tableaux are standard, we obtain Fig. 4.

Each route in this figure results in a unique standard tableau. Conversely, each standard tableau uniquely defines its predecessors along the route. This follows from the fact that removal of the highest number from a standard tableau for N numbers produces a standard tableau for $(N - 1)$

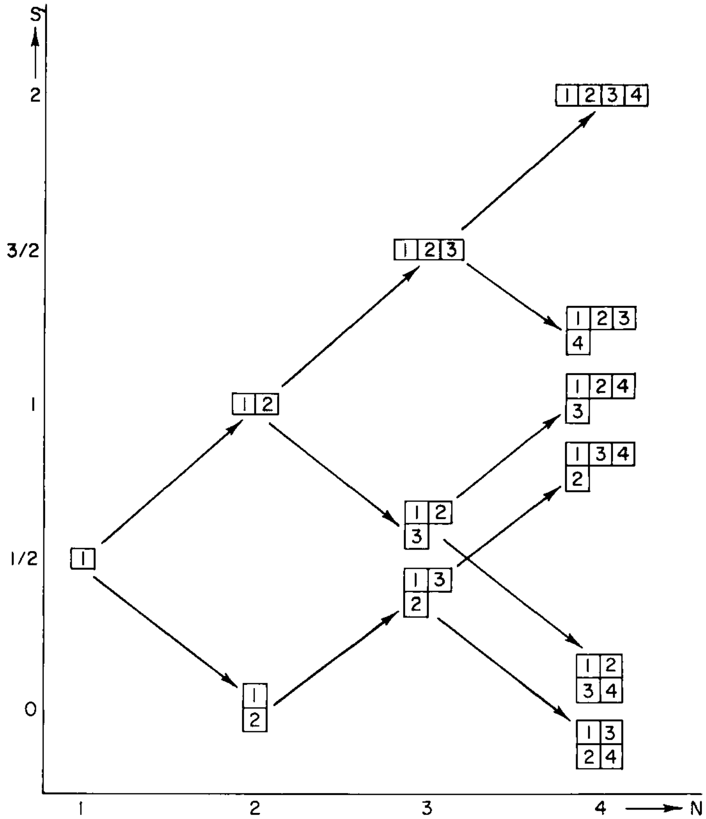
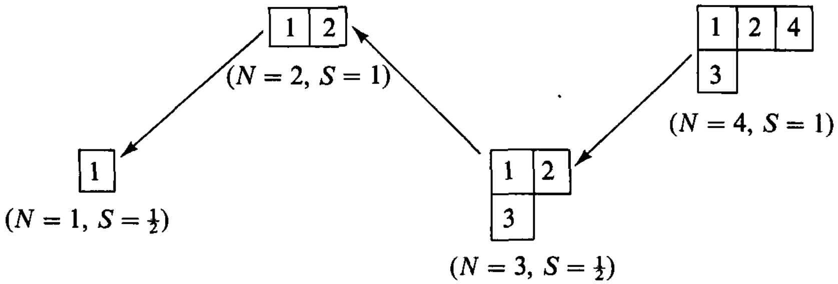


Fig. 4. YK branching diagram for standard Young tableaux.

numbers. Thus, for example, one can work backward from $\begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3&\multicolumn{2}{c}{}\\\cline{1-1}\end{array}$ in the following way:



The significance of this is that each standard tableau can be uniquely associated with a YK branching route, and therefore can be uniquely associated with a YK spin function. To use the example of $N = 4$, $S = 1$ (or $D = \begin{array}{|c|c|c|}\hline \ & \ & \ \\\hline \ &\multicolumn{2}{c}{}\\\cline{1-1}\end{array}$), we have the correspondence

Standard tableau	Branching route
$\begin{array}{\|c\|c\|c\|}\hline 1&2&3\\\hline 4&\multicolumn{2}{c}{}\\\cline{1-1}\end{array}$	
$\begin{array}{\|c\|c\|c\|}\hline 1&2&4\\\hline 3&\multicolumn{2}{c}{}\\\cline{1-1}\end{array}$	
$\begin{array}{\|c\|c\|c\|}\hline 1&3&4\\\hline 2&\multicolumn{2}{c}{}\\\cline{1-1}\end{array}$	

It will be observed that each number on the upper row of a standard tableau corresponds to an upward movement in the associated branching route, and each number on the lower row corresponds to a downward movement.

Strictly speaking, it is not the tableau itself that corresponds to a YK branching route, but the unique "chain" of tableau predecessors from which it derives. For example, the branching route  is a shorthand for the tableau chain

$$\begin{array}{|c|}\hline 1\\\hline\end{array} \to \begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array} \to \begin{array}{|c|c|}\hline 1&3\\\hline 2&\multicolumn{1}{c}{}\\\cline{1-1}\end{array} \to \begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&\multicolumn{2}{c}{}\\\cline{1-1}\end{array}$$

Such a chain involves the addition of one number at a time, and will be called a 1-*chain*.

In general, we denote by $T_r^{D,k}$ the standard tableau obtained from T_r^D by removing its k highest numbers, *viz.*, $N, N-1, \ldots, N-k+1$. Thus the 1-chain defined by T_r^D is written

$$T_r^{D,N-1} \rightarrow T_r^{D,N-2} \rightarrow \cdots \rightarrow T_r^{D,1} \rightarrow T_r^D.$$

Each standard tableau is also associated with a unique 2-*chain*, if N is even. Removal of two numbers from a standard tableau results in a smaller tableau which is also standard. Thus one can work backward from a given standard tableau and define its predecessors in a Serber-type genealogical scheme, for example,

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1} \end{array} \rightarrow \begin{array}{|c|c|c|c|}\hline 1 & 2 & 4 & 6 \\ \hline 3 & 5 \\ \cline{1-2} \end{array}.$$

In general,

$$T_r^{D,N-2} \rightarrow T_r^{D,N-4} \rightarrow \cdots \rightarrow T_r^{D,2} \rightarrow T_r^D.$$

In other words, standard tableaux can be considered constructed according to the Serber branching diagram of Fig. 5. We have indicated in each case the pair of numbers being added, and their positions relative to the original tableau.

At each stage in such a branching diagram, a geminal pair of numbers $2\mu-1, 2\mu$ is added to a tableau containing $\mu-1$ geminal pairs. It will be observed that the addition of $\cdots \begin{array}{|c|c|}\hline 2\mu-1 & 2\mu \\ \hline \end{array}$ always corresponds to $s_\mu = 1$, and the addition of two numbers on the same column always has the effect of adding $s_\mu = 0$. There is an ambiguity, however, when $2\mu-1$ and 2μ are on neither the same row nor the same column. One case must correspond to the addition of $s_\mu = 1$ and the other to $s_\mu = 0$. We are free to make a choice, so long as it is consistent. In the following pages, we shall associate

$$\begin{array}{rl} \cdots & \boxed{2\mu} \\ \boxed{2\mu-1} & \end{array} \qquad \text{with } s_\mu = 1$$

and

$$\begin{array}{ll} \cdots & \boxed{2\mu-1} \\ \boxed{2\mu} & \end{array} \qquad \text{with } s_\mu = 0.$$

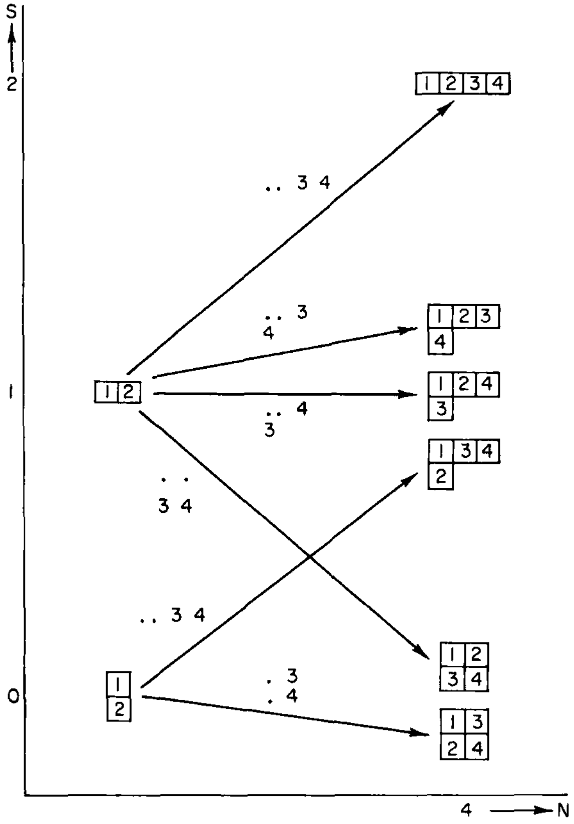


Fig. 5. Serber branching diagram for standard Young tableaux.

Now it is clear that the concept of tableau chains provides the link between Young's theory of the symmetric group and the genealogical construction of spin functions. However, we have already pointed out that Young operators do not generate YK or Serber spin functions. Clearly, this is because they do not, in themelves, carry information specific to 1-chains or 2-chains.

We begin to remedy this deficiency by defining *chains of Young idempotents*. Suppose that E_r^D and $E_r^{D,k}$ are the Young idempotents for the tableaux T_r^D and $T_r^{D,k}$, respectively. Then the m-chain of standard tableaux

$$T_r^{D,N-m} \to T_r^{D,N-2m} \to \cdots \to T_r^{D,m} \to T_r^D$$

is associated with the m-chain

$$E_r^{D,N-m} \to E_r^{D,N-2m} \to \cdots \to E_r^{D,m} \to E_r^{D}$$

of Young idempotents. (It is assumed that N is a multiple of m.)

Carrying this one step further, we define $L_r^{D,k}$ to be the minimal left ideal generated by the Young idempotent $E_r^{D,k}$. Thus each standard tableau T_r^D defines a unique m-chain

$$L_r^{D,N-m} \to L_r^{D,N-2m} \to \cdots \to L_r^{D,m} \to L_r^{D}$$

of minimal left ideals.

B. Chains of Young Idempotents and Genealogical Spin Functions: A Heuristic Argument

It was mentioned previously that the YK spin functions for fixed N and S form a basis for a special orthogonal irrep of S_N in which the representations of $S_{N-1}, S_{N-2}, \ldots, S_1$ are also fully reduced. The representation is said to be *adapted to the sequence of groups*

$$S_N \supset S_{N-1} \supset S_{N-2} \supset \cdots \supset S_1.$$

We shall say that a representation with this property is *YK-adapted.*

In a similar way, the Serber functions (for even N) are adapted to the sequence

$$S_N \supset S_{N-2} \supset S_{N-4} \supset \cdots \supset S_2.$$

In addition, every geminal two-electron subgroup of S_N is represented irreducibly. A representation with these two properties is said to be *Serber-adapted.*

The adaptation of representations to sequences of nested symmetric groups is the group-theoretical significance of a genealogical spin-coupling scheme. Genealogical spin eigenfunctions of the types we seek will be generated by operators that form YK- and Serber-adapted bases for minimal left ideals of the symmetric group algebra. But what does this mean?

Suppose that $L(Y)$ is a subspace of the group algebra $A(S_N)$, spanned by a basis $B(Y)$ and having the following properties:

(1) $L(Y)$ is invariant under left multiplications by elements of S_N and is equivalent to the minimal left ideal L_r^D with respect to these left multiplications.[2]

---

[2] The meaning of "equivalence" as applied to minimal left ideals is discussed in Boerner (1963, p. 56; see also Weyl, 1950, p. 283, 294).

(2) $B(Y)$ can be partitioned into subsets, each of which transforms internally under left multiplications by elements of S_{N-1}, and each of which spans a carrier space for an *irrep* of S_{N-1}. Each of these subsets can in turn be partitioned into subsets, each of which spans a carrier space for an irrep of S_{N-2}. And so on.

From property (1), $B(Y)$ spans a carrier space for an irreducible representation of S_N, i.e., it spans a minimal left ideal in $A(S_N)$. Property (2) is what is meant by saying that $B(Y)$ is adapted to the sequence $S_N \supset S_{N-1} \supset S_{N-2} \supset \cdots \supset S_1$. Clearly, $L(Y)$ is a carrier space for a YK-adapted representation of S_N.

In a similar way, a subspace $L(S)$ of $A(S_N)$ is a carrier space for a Serber-adapted representation if:

(1) Its basis $B(S)$ transforms internally under left multiplications by elements of S_N and is equivalent to a basis for the minimal left ideal L_r^D with respect to these left multiplications.

(2) $B(S)$ can be partitioned into subsets that transform internally under left multiplications by elements of S_{N-2} and which span irreps of S_{N-2}; these subsets can be partitioned into bases for irreps of S_{N-4}; etc.

(3) The elements of $B(S)$ are either symmetric or antisymmetric with respect to left multiplications by geminal transpositions $(2\mu - 1, 2\mu)$.

Before defining orthogonal matric bases for genealogical representations, it is instructive to see what predictions can be made about the structure of such operators by extending the present argument. We shall see that idempotent generators for YK- and Serber-adapted carrier spaces can be deduced rather easily.

The minimal left ideal associated with the standard tableau T_r^D is defined to be $L_r^D = \{XE_r^D\}$, where X sweeps the whole group algebra. It can be shown (Rutherford, 1948, pp. 20 ff.) that Young idempotents have the property

$$E_r^D X E_r^D = \theta^D \cdot i[XE_r^D] E_r^D,$$

where $\theta^D = (N!/d^D) > 0$ does not depend on r, and $i[XE_r^D]$ is the coefficient of the identity in XE_r^D, when it is expanded in terms of the group elements. It follows that

$$(XE_r^D)(XE_r^D) = \theta^D \cdot i[XE_r^D](XE_r^D),$$

so that (XE_r^D) is essentially idempotent if it contains the identity. In other words, *new idempotent generators in L_r^D can be made by left multiplying E_r^D.*

Consider, for example, the element

$$E_Y(D, r) = E_r^{D,N-1} E_r^{D,N-2} \cdots E_r^{D,1} E_r^D.$$

This operator belongs to L_r^D. To the left, it has $E_r^{D,N-1}$, which generates $L_r^{D,N-1}$; $(E_r^{D,N-1}E_r^{D,N-2})$, belonging to $L_r^{D,N-2}$; $(E_r^{D,N-1}E_r^{D,N-2}E_r^{D,N-3})$, belonging to $L_r^{D,N-3}$; etc. Thus $E_Y(D, r)$ behaves under left multiplications by elements of S_{N-k} (where $k = 1, 2, \ldots, N-1$) like an element of $L_r^{D,k}$. The Young idempotent $E_r{}^D$ has been "YK-adapted" by multiplying it from the left by the 1-chain

$$E_1^{D,N-1} \to E_r^{D,N-2} \to \cdots \to E_r^{D,1}$$

of Young idempotents from which it derives.

Similarly, we may expect a Serber-adapted idempotent to take the form

$$E_S(D, r) = S_r^{D,N-2}E_r^{D,N-2}S_r^{D,N-4}E_r^{D,N-4} \cdots S_r^{D,2}E_r^{D,2}S_r{}^D E_r{}^D,$$

where $S_r^{D,2k}$ either symmetrizes or antisymmetrizes the geminal pair $(N-2k-1, N-2k)$. Since the operators to the left of $S_r^{D,2k}$ do not contain the electron labels on which it operates, the pair-symmetry operators can all be brought out to the left:

$$E_S(D, r) = (S_r^{D,N-2}S_r^{D,N-4} \cdots S_r^{D,2}S_r{}^D) \cdot (E_r^{D,N-2}E_r^{D,N-4} \cdots E_r^{D,2}E_r{}^D).$$

Thus, when $E_S(D, r)$ is applied to a primitive function, it will generate a function which is either symmetric or antisymmetric in each geminal pair.

Assuming that $E_Y(D, r)$ and $E_S(D, r)$, when expanded in terms of the group elements, contain the identity, they are essentially idempotent. However, they are not Hermitian, so they cannot be the idempotent diagonal elements of the matric bases we seek.

It is easy to see that the following operators *are* Hermitian:

$$E_Y(D, r)E_Y{}^\dagger(D, r) = E_r^{D,N-1} \cdots E_r^{D,1}E_r{}^D E_r^{D\dagger}E_r^{D,1\dagger} \cdots E_r^{D,N-1\dagger},$$

$$E_S(D, r)E_S{}^\dagger(D, r) = G_r{}^D E_r^{D,N-2} \cdots E_r^{D,2}E_r{}^D E_r^{D\dagger}E_r^{D,2\dagger} \cdots E_r^{D,N-2\dagger}G_r{}^D,$$

in which $G_r{}^D = (S_r^{D,N-2} \cdots S_r^{D,2}S_r{}^D) = G_r^{D\dagger}$.

It can be shown that these operators are, in fact, Hermitian idempotents generating YK- and Serber-adapted carrier spaces for irreducible representations of S_N. It can also be shown, however, that they do not multiply like the diagonal elements of a matric basis. It may be that

$$[E_Y(D, r)E_Y{}^\dagger(D, r)][E_Y(D, s)E_Y{}^\dagger(D, s)] \neq 0,$$

for example. Thus these operators cannot be used to generate orthogonal basis functions.

We present in the next section matric bases for YK- and Serber-adapted orthogonal representations. It will be seen that these matric bases are symmetry adapted in a way similar to $E_Y E_Y{}^\dagger$ and $E_S E_S{}^\dagger$. Their definitions differ only as required to obtain the correct multiplication properties.

C. Matric Bases for Orthogonal YK-Adapted Representations

The standard tableau $T_r{}^D$ defines the 1-chain

$$T_r^{D,N-1} \to T_r^{D,N-2} \to \cdots \to T_r^{D,1} \to T_r{}^D$$

of standard tableaux, where $T_r^{D,N-1} = \boxed{1}$ for every D and r.

We define for this 1-chain a chain of idempotent operators in the following manner:

$$\begin{aligned} e_r^{D,N-1} &= I, \qquad \text{the identity}, \\ e_r^{D,N-2} &= (E_r^{D,N-2} e_r^{D,N-1})^\dagger (E_r^{D,N-2} e_r^{D,N-1})/k_r^{D,N-2}, \\ &\ldots, \\ e_r^{D,i} &= (E_r^{D,i} e_r^{D,i+1})^\dagger (E_r^{D,i} e_r^{D,i+1})/k_r^{D,i}, \\ &\ldots, \\ e_r{}^D &= (E_r{}^D e_r^{D,1})^\dagger (E_r{}^D e_r^{D,1})/k_r{}^D, \end{aligned} \tag{18}$$

where $k_r{}^D$ is the number

$$k_r{}^D = o_{\mathscr{R}}{}^D \cdot \theta^D \cdot \rho_r{}^D,$$

in which $\rho_r{}^D = i[C_r{}^D E_r{}^D e_r^{D,1}]$ and $o_{\mathscr{R}}{}^D$ = order of row group for any tableau belonging to diagram D. It should be noted that these operators are Hermitian.

The idempotents $e_r{}^D$ are used to construct the matric basis elements

$$e_{rs}^D = (E_r{}^D e_r^{D,1})^\dagger p_{rs}^D (E_s{}^D e_s^{D,1})/(k_r{}^D k_s{}^D)^{1/2}. \tag{19}$$

The diagonal elements e_{rr}^D of this basis are identical to the idempotents $e_r{}^D$ defined by Eq. (18).

For application to primitive functions, it is more convenient to use an alternative expression for the matric basis:

$$\begin{aligned} e_{rs}^D &= e_r^{D,1} C_r{}^D R_r{}^D p_{rs}^D R_s{}^D C_s{}^D e_s^{D,1}/(k_r{}^D k_s{}^D)^{1/2} \\ &= e_r^{D,1} p_{rs}^D C_s{}^D R_s{}^D R_s{}^D C_s{}^D e_s^{D,1}/[\theta^D o_{\mathscr{R}}{}^D (\rho_r{}^D \rho_s{}^D)^{1/2}] \\ &= e_r^{D,1} p_{rs}^D C_s{}^D R_s{}^D C_s{}^D e_s^{D,1}/[\theta^D (\rho_r{}^D \rho_s{}^D)^{1/2}], \end{aligned}$$

or

$$e_{rs}^D = e_r^{D,1} p_{rs}^D C_s{}^D E_s{}^D e_s^{D,1}/[\theta^D (\rho_r{}^D \rho_s{}^D)^{1/2}]. \tag{20}$$

These equations define operators built around Young operators, but adapted to the genealogy of 1-chains through Eqs. (18).

The definitions are most easily understood by working an example. Let $D = \begin{array}{|c|c|}\hline \ & \ \\ \hline \ \\ \cline{1-1}\end{array}$, for which the standard tableaux are

$$T_1{}^D = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \text{and} \quad T_2{}^D = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array},$$

so that $p_{12}^D = (2, 3) = p_{21}^D$.

The 1-chain defined by $T_1{}^D$ is

$$T_1^{D,2} = \begin{array}{|c|}\hline 1 \\ \hline\end{array} \to T_1^{D,1} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array} \to T_1{}^D = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array},$$

for which the Young idempotents are

$$E_1^{D,2} = I \to E_1^{D,1} = \mathscr{S}_{12} \to E_1{}^D = \mathscr{S}_{12}\mathscr{A}_{13}.$$

Neglecting numerical factors,

$$\begin{aligned} e_1^{D,2} &= I, \\ e_1^{D,1} &= e_1^{D,2}C_1^{D,1}E_1^{D,1}e_1^{D,2} = I \cdot I \cdot \mathscr{S}_{12} \cdot I = \mathscr{S}_{12}, \\ e_1{}^D &= e_1^{D,1}C_1{}^DE_1{}^De_1^{D,1} = \mathscr{S}_{12}\mathscr{A}_{13}\mathscr{S}_{12}\mathscr{A}_{13}\mathscr{S}_{12}. \end{aligned}$$

The 1-chain defined by $T_2{}^D$ is

$$T_2^{D,2} = \begin{array}{|c|}\hline 1 \\ \hline\end{array} \to T_2^{D,1} = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline\end{array} \to T_2{}^D = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array},$$

for which the Young idempotents are

$$E_2^{D,2} = I \to E_2^{D,1} = \mathscr{A}_{12} \to E_2{}^D = \mathscr{S}_{13}\mathscr{A}_{12}.$$

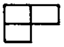
Neglecting numerical factors,

$$\begin{aligned} e_2^{D,2} &\overset{d}{=} I, \\ e_2^{D,1} &= e_2^{D,2}C_2^{D,1}E_2^{D,1}e_2^{D,2} = I \cdot \mathscr{A}_{12}\mathscr{A}_{12} \cdot I = 2\mathscr{A}_{12}, \\ e_2{}^D &= e_2^{D,1}C_2{}^DE_2{}^De_2^{D,1} = 4\mathscr{A}_{12}\mathscr{A}_{12}\mathscr{S}_{13}\mathscr{A}_{12}\mathscr{A}_{12} = 16\mathscr{A}_{12}\mathscr{S}_{13}\mathscr{A}_{12}. \end{aligned}$$

The entire matric basis, then, consists of the operators

$$\begin{aligned}
e^D_{11} &= e_1{}^D = \mathscr{S}_{12}\mathscr{A}_{13}\mathscr{S}_{12}\mathscr{A}_{13}\mathscr{S}_{12},\\
e^D_{21} &= e_2^{D,1}p^D_{21}C_1{}^DE_1{}^De_1^{D,1} = 2\mathscr{A}_{12}\cdot(2,3)\cdot\mathscr{A}_{13}\mathscr{S}_{12}\mathscr{A}_{13}\mathscr{S}_{12},\\
e^D_{12} &= e_1^{D,1}p^D_{12}C_2{}^DE_2{}^De_2^{D,1} = 2\mathscr{S}_{12}\cdot(2,3)\cdot\mathscr{A}_{12}\mathscr{S}_{13}\mathscr{A}_{12}\mathscr{A}_{12}\\
&= 4\mathscr{S}_{12}\cdot(2,3)\cdot\mathscr{A}_{12}\mathscr{S}_{13}\mathscr{A}_{12},\\
e^D_{22} &= e_2{}^D = 16\mathscr{A}_{12}\mathscr{S}_{13}\mathscr{A}_{12}.
\end{aligned}$$

The whole matric basis is not required for the construction of basis functions for the irrep. The operators $\{e^D_{11}, e^D_{21}\}$ span a minimal left ideal associated with $T_1{}^D$; similarly, $\{e^D_{12}, e^D_{22}\}$ span a minimal left ideal associated with $T_2{}^D$. Either of these subsets can be used to generate basis functions.

As an example, we apply e^D_{12} and e^D_{22} to the spin product function $\theta = \alpha\beta\alpha$. Since the diagram $D =$  corresponds to $S = \frac{1}{2}$, we should obtain YK spin functions for $N = 3$, $S = M = \frac{1}{2}$. We have

$$e^D_{12} = 4\mathscr{S}_{12}\cdot(2,3)\cdot\mathscr{P},$$

where $\mathscr{P} = \mathscr{A}_{12}\mathscr{S}_{13}\mathscr{A}_{12}$. Thus

$$\begin{aligned}
\mathscr{P}\theta &= \mathscr{A}_{12}\mathscr{S}_{13}(\alpha\beta\alpha - \beta\alpha\alpha) = \mathscr{A}_{12}(2\alpha\beta\alpha - \beta\alpha\alpha - \alpha\alpha\beta)\\
&= (2\alpha\beta\alpha - 2\beta\alpha\alpha - \beta\alpha\alpha + \alpha\beta\alpha) = 3(\alpha\beta\alpha - \beta\alpha\alpha),
\end{aligned}$$

so

$$\begin{aligned}
e^D_{12}\theta &= 12\mathscr{S}_{12}\cdot(2,3)\cdot(\alpha\beta\alpha - \beta\alpha\alpha) = 12\mathscr{S}_{12}(\alpha\alpha\beta - \beta\alpha\alpha)\\
&= 12(2\alpha\alpha\beta - \beta\alpha\alpha - \alpha\beta\alpha) = 12[2\alpha\alpha\beta - (\alpha\beta + \beta\alpha)\alpha]
\end{aligned}$$

and

$$e^D_{22}\theta = 16\mathscr{P}\theta = 48(\alpha\beta - \beta\alpha)\alpha.$$

These are, indeed, the (unnormalized) YK spin functions obtained for $N = 3$, $S = \frac{1}{2}$, $M = \frac{1}{2}$ from the spin-coupling Eqs. (2). Notice that $e^D_{12}\theta$ corresponds to the branching route , while $e^D_{22}\theta$ corresponds to , and that these functions are orthogonal.

The same functions, within a numerical factor, are obtained by means of the matric basis elements e^D_{11} and e^D_{21}.

D. Matric Bases for Orthogonal Serber-Adapted Representations

When N is even, the standard tableau T_r^D defines the 2-chain

$$T_r^{D,N-2} \to T_r^{D,N-4} \to \cdots \to T_r^{D,2} \to T_r^D,$$

where $T_r^{D,N-2}$ is either $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline \end{array}$ or $\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline \end{array}$, depending on T_r^D.

For T_r^D, a geminal operator S_r^D is defined in terms of the positions of the two highest numbers, $N-1$ and N. Denoting the row and column on which a number k appears as r_k and c_k, we define

$$S_r^D = \begin{cases} [I + (N-1, N)]/2 & \text{if} \quad r_{N-1} = r_N \quad \text{or} \quad r_{N-1} > r_N, \\ [I - (N-1, N)]/2 & \text{if} \quad c_{N-1} = c_N \quad \text{or} \quad (r_{N-1} < r_N, c_{N-1} \neq c_N). \end{cases}$$

In other words, S_r^D symmetrizes the numbers $(N-1)$ and N if T_r^D contains these numbers in the positions

$$\cdots \begin{array}{|c|c|}\hline (N-1) & N \\ \hline \end{array} \qquad \text{or} \qquad \begin{array}{ccc} & \cdots & \boxed{N} \\ \boxed{N-1} & & \end{array},$$

but antisymmetrizes them if T_r^D contains

$$\begin{array}{ccc} & \cdots & \boxed{N-1} \\ \boxed{N} & & \end{array} \qquad \text{or} \qquad \begin{array}{|c|}\hline N-1 \\ \hline N \\ \hline \end{array}.$$

Geminal operators $S_r^{D,2k}$ for other tableaux $T_r^{D,2k}$ in the 2-chain are defined analogously.

A set of Hermitian idempotents is defined recursively for the 2-chain:

$$e_r^{D,N-2} = S_r^{D,N-2},$$

$$e_r^{D,N-4} = (E_r^{D,N-4} S_r^{D,N-4} e_r^{D,N-2})^\dagger (E_r^{D,N-4} S_r^{D,N-4} e_r^{D,N-2})/k_r^{D,N-4},$$

$$\cdots,$$

$$e_r^D = (E_r^D S_r^D e_r^{D,2})^\dagger (E_r^D S_r^D e_r^{D,2})/k_r^D, \tag{21}$$

where k_r^D is the number $k_r^D = o_{\mathscr{R}}^D \cdot \theta^D \cdot \rho_r^D$ in which $\rho_r^D = i[C_r^D E_r^D S_r^D e_r^{D,2}]$ and $o_{\mathscr{R}}^D =$ order of the row group for any tableau belonging to diagram D.

The idempotents e_r^D are used to construct the matric basis elements

$$e_{rs}^D = (E_r^D S_r^D e_r^{D,2})^\dagger p_{rs}^D (E_s^D S_s^D e_s^{D,2})/(k_r^D k_s^D)^{1/2}. \tag{22}$$

It should be noted that a diagonal element e_{rr}^{D} in this basis is identical to the element $e_r{}^D$ defined by (21).

As in the previous case, the matric basis elements can be given in a slightly simpler form. The result is

$$e_{rs}^{D} = e_r^{D,2} S_r{}^D p_{rs}^{D} C_s{}^D E_s{}^D S_s{}^D e_s^{D,2} / [\theta^D (\rho_r{}^D \rho_s{}^D)^{1/2}]. \tag{23}$$

As an example of the application of these operators, we generate the Serber spin functions for $N = 4$, $S = 1$, $M = 0$, using the primitive function $\alpha\beta\alpha\beta$. The Young diagram is $D = \begin{array}{|c|c|c|}\hline \ & \ & \ \\ \hline \ & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}$, for which the standard tableaux are

$$T_1{}^D = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}, \qquad T_2{}^D = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}, \qquad T_3{}^D = \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array},$$

so that

$$p_{11}^{D} = I, \qquad p_{21}^{D} = (3, 4), \qquad p_{31}^{D} = (2, 3)(3, 4),$$

and

$$E_1{}^D = \mathscr{S}_{123}\mathscr{A}_{14}, \qquad E_2{}^D = \mathscr{S}_{124}\mathscr{A}_{13}, \qquad E_3{}^D = \mathscr{S}_{134}\mathscr{A}_{12}.$$

The 2-chain defined by $T_1{}^D$ is

$$T_1^{D,2} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array} \rightarrow T_1{}^D = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}.$$

Thus

$$e_1^{D,2} = e_{11}^{D,2} = S_1^{D,2} = \mathscr{S}_{12} \quad \text{and} \quad S_1{}^D = \mathscr{A}_{34}.$$

The 2-chain defined by $T_2{}^D$ is

$$T_2^{D,2} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array} \rightarrow T_2{}^D = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}.$$

Thus

$$e_2^{D,2} = e_{22}^{D,2} = S_2^{D,2} = \mathscr{S}_{12} \quad \text{and} \quad S_2{}^D = \mathscr{S}_{34}.$$

The 2-chain defined by $T_3{}^D$ is

$$T_3^{D,2} = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline \end{array} \to T_3{}^D = \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 \\ \cline{1-1} \end{array},$$

so that

$$e_3^{D,2} = e_{33}^{D,2} = S_3^{D,2} = \mathscr{A}_{12} \quad \text{and} \quad S_3{}^D = \mathscr{S}_{34}.$$

The matric basis elements $\{e_{k1}^D \mid k = 1, 2, 3\}$, which span a minimal left ideal associated with $T_1{}^D$, are therefore [using Eq. (23) and neglecting numerical factors]

$$\begin{aligned} e_{11}^D &= \mathscr{S}_{12}\mathscr{A}_{34}\mathscr{P}, \\ e_{21}^D &= \mathscr{S}_{12}\mathscr{S}_{34} \cdot (3, 4) \cdot \mathscr{P}, \\ e_{31}^D &= \mathscr{A}_{12}\mathscr{S}_{34} \cdot (2, 3)(3, 4) \cdot \mathscr{P}, \end{aligned}$$

where $\mathscr{P} = \mathscr{A}_{14}\mathscr{S}_{123}\mathscr{A}_{14}\mathscr{A}_{34}\mathscr{S}_{12}$.

Applying these operators to $\theta = \alpha\beta\alpha\beta$, one obtains the Serber functions shown previously. The branching routes can be read directly from the geminal symmetrizers and antisymmetrizers in the matric basis elements.

E. General Definition of the Genealogical Matric Bases

It is convenient to treat the matric bases for 1-chains and 2-chains together under one master formula. Let the m-chain defined by the standard tableau $T_r{}^D$ be denoted by

$$T_r^{D,N-m} \to T_r^{D,N-2m} \to \cdots \to T_r^{D,m} \to T_r{}^D,$$

where m is a factor of N.

For each standard tableau $T_r^{D,jm}$ in this chain, a Hermitian operator $M_r^{D,jm}$ is defined in terms of only the highest m numbers, i.e., the numbers $N - jm$, $N - jm - 1, \ldots N - (j + 1)m + 1$. When $m = 1$, this operator is taken to be the identity. When $m = 2$, it is defined to be a two-electron symmetrizer or antisymmetrizer, as discussed previously.

A set of Hermitian idempotents is defined recursively in terms of each m-chain:

$$\begin{aligned} e_r^{D,N-m} &= M_r^{D,N-m}, \\ e_r^{D,N-2m} &= (E_r^{D,N-2m} M_r^{D,N-2m} e_r^{D,N-m})^\dagger \\ &\quad \cdot (E_r^{D,N-2m} M_r^{D,N-2m} e_r^{D,N-m}) / k_r^{D,N-2m}, \\ &\ldots, \\ e_r{}^D &= (E_r{}^D M_r{}^D e_r^{D,m})^\dagger (E_r{}^D M_r{}^D e_r^{D,m}) / k_r{}^D, \end{aligned}$$

where ${k_r}^D = {o_{\mathscr{R}}}^D \cdot \theta^D \cdot {\rho_r}^D$, in which ${\rho_r}^D = i[{C_r}^D {E_r}^D {M_r}^D e_r^{D,m}]$ and ${o_{\mathscr{R}}}^D =$ order of row group for any tableau belonging to diagram D.

These idempotents are used to define the matric basis

$$e_{rs}^D = ({E_r}^D {M_r}^D e_r^{D,m})^\dagger p_{rs}^D ({E_s}^D {M_s}^D e_s^{D,m}) / ({k_r}^D {k_s}^D)^{1/2}, \tag{24}$$

in which, it will be noted, $e_{rr}^D = {e_r}^D$.

It is convenient to use the matric basis elements in the simpler form

$$e_{rs}^D = e_r^{D,m} {M_r}^D p_{rs}^D {C_s}^D {E_s}^D {M_s}^D e_s^{D,m} / [\theta^D ({\rho_r}^D {\rho_s}^D)^{1/2}]. \tag{25}$$

For use in generating basis functions for the irrep of S_N labeled by D, a subset $\{e_{rs}^D \mid s \text{ fixed}\}$ of the matric basis is used. The operators in this subset all have the form

$$e_{rs}^D = (\text{number}) \cdot e_r^{D,m} {M_r}^D p_{rs}^D {\mathscr{P}_s}^D, \tag{26a}$$

where

$${\mathscr{P}_s}^D = {C_s}^D {E_s}^D {M_s}^D e_s^{D,m} \tag{26b}$$

is *fixed*.

F. Discussion

We shall prove in the next section that the matric bases defined by Eqs. (18)–(25) can be used to generate basis functions for orthogonal representations of S_N. More precisely, we will show that (1) none of the elements e_{rs}^D vanishes or "blows up", (2) these elements multiply like a matric basis, (3) they possess the adjoint property $e_{rs}^{D\dagger} = e_{sr}^D$, (4) they are linearly independent and span the group algebra $A(S_N)$, and (5) the diagonal elements e_{rr}^D satisfy Eq. (14).

That the matric bases are YK-adapted (when $m = 1$) or Serber-adapted (when $m = 2$) is easier to see. Using Eq. (25), neglecting numerical factors, and noting that ${M_r}^D$ commutes with $e_r^{D,m}$, $e^{D,2m}$, etc., we find

$$\begin{aligned} e_{rs}^D &= e_r^{D,m} {M_r}^D p_{rs}^D {C_s}^D {E_s}^D {M_s}^D e_s^{D,m} \\ &= {M_r}^D e_r^{D,m} {C_r}^D {E_r}^D p_{rs}^D e_s^{D,m} {M_s}^D \\ &= {M_r}^D (e_r^{D,2m} M_r^{D,m} C_r^{D,m} E_r^{D,m} M_r^{D,m} e_r^{D,2m}) \\ &\quad \cdot {C_r}^D {E_r}^D p_{rs}^D \cdot (e_s^{D,2m} M_s^{D,m} C_s^{D,m} E_s^{D,m} M_s^{D,m} e_s^{D,2m}) {M_s}^D \\ &= {M_r}^D M_r^{D,m} (e_r^{D,2m} C_r^{D,m} E_r^{D,m} M_r^{D,m} e_r^{D,2m}) {C_r}^D {E_r}^D p_{rs}^D \cdots \\ &= \cdots \\ &= {G_r}^D E_r^{D,N-m} \cdots E_r^{D,N-2m} \cdots E_r^{D,m} \cdots {E_r}^D \cdots \end{aligned} \tag{27}$$

where $G_r^D = M_r^D M_r^{D,m} \cdots M_r^{D,N-m}$ is a product of commuting operators. When $m = 1$, G_r^D is simply the identity. When $m = 2$, it is a string of geminal symmetrizers and antisymmetrizers.

Comparison of Eq. (27) with the heuristically derived operators $E_Y E_Y^\dagger$ and $E_S E_S^\dagger$ of the previous section shows that e_{rs}^D is YK-adapted when $m = 1$ and Serber-adapted when $m = 2$.

Orthogonal YK-adapted representation matrices were first obtained by Young (1932). This is the representation known in the literature as "Young's orthogonal representation." Pauncz (1967) has shown that this representation is identical to that obtained by Yamanouchi. A matric basis for such a representation can be obtained from the relations (5) between orthogonal matric basis elements and permutations. One obtains the so-called "orthogonal units" (Rutherford, 1948, p. 50)

$$o_{rs}^D = \frac{d^D}{N!} \sum_P [P]_{rs}^D P, \tag{28}$$

where the sum runs over the *entire* symmetric group. Goddard (1967, 1968) has employed this matric basis in quantum-chemical calculations.

In nuclear theory, Jahn and co-workers (Jahn and van Wieringen, 1951; Elliott *et al.*, 1953; Jahn, 1954) have used matric bases for orthogonal YK- *and* Serber-adapted representations. The latter were obtained from the orthogonal units (28) by finding the transformation between YK and Serber representations.

General discussions of matric bases, considered according to their expansions in permutations, have been given by Matsen and co-workers (Matsen, 1964; Matsen *et al.*, 1966; Klein *et al.*, 1970).

In all of these accounts, matric basis elements were described as linear combinations of all $N!$ permutations in S_N. Thus matric bases were expressed as sets of Wigner operators. The disadvantages of this approach were discussed earlier.

To the author's knowledge, the only previous attempt to obtain matric bases directly from the standard Young tableaux was the derivation by Thrall (1941) of the "seminormal units." These have been discussed by Rutherford (1948, Ch. 3). The work reported in the present paper can be viewed as an extension of Thrall's approach to *orthogonal* representations useful in quantum-mechanical applications. The matric bases defined here generate irreps equivalent to the YK and Serber representations. It is hoped that the operator properties proved in Section V will lead to closed expressions for the Serber-adapted representation matrices similar to those obtained by Young for the YK matrices. This will be the subject of future work.

The formulas given in the previous section would appear to avoid the drawbacks of other methods for obtaining basis functions. Referring to Eq. (26), one sees that basis functions for *any* of the distinct irreps of S_N can be generated by a set of operators constructed from symmetrizers, antisymmetrizers, and the permutations p_{rs}^D relating standard tableaux. Furthermore, the "right half" of each operator, given by Eq. (26b), is fixed throughout the calculation.

Although the matric bases presented here are defined recursively, this does not cause serious computational difficulties. The recursion gives rise to a number of row and column operators which must be applied in succession to a primitive function. As can be seen from the examples in the last section, one applies a symmetrizer or antisymmetrizer to the primitive, collects terms, and then applies another. The operators are all "read" directly from the standard tableaux. A computer program for such a procedure would not require large amounts of storage—the chief drawback of other approaches. Such a program *would* have to perform very many permutations and collections of terms, but these operations involve only data transferrals and integer arithmetic and can be performed quickly. A computer program is being written to generate Serber-type spin functions by means of the matric basis elements (23).

It should be noted that YK- and Serber-adapted matric bases have been defined here for every Young diagram, not just those corresponding to spin representations. Thus operator bases have been defined for space representations also, and complete wavefunctions can be constructed conveniently. This subject is discussed in Section VI.

V. Proofs for Section IV

A. Basic Lemmas

Before proceeding to the lemmas and theorems specific to orthogonal matric bases, we summarize some elementary results that will be needed.

The definitions (18)–(25) used in the construction of matric bases involve numerical factors $i[x]$, the coefficient of the identity in an element x of the group algebra. This function defined on $A(S_N)$ has two properties which we shall find useful.

LEMMA 1. *If μ, ν are numbers and x, y are elements of $A(S_N)$, then* $i[\mu x + \nu y] = \mu \cdot i[x] + \nu \cdot i[y]$.

Proof. If $x = \sum \xi(P)P$ and $y = \sum \eta(P)P$, then

$$i[\mu x + \nu y] = i\{\sum [\mu\xi(P) + \nu\eta(P)]P\} = \mu\xi(I) + \nu\eta(I).$$

But $\mu \cdot i[x] + \nu \cdot i[y] = \mu\xi(I) + \nu\eta(I)$ also.

LEMMA 2. *If x and y are elements of $A(S_N)$, then $i[xy] = i[yx]$.*

Proof. If $x = \sum \xi(P)P$ and $y = \sum \eta(P')P'$, then

$$i[xy] = \sum \xi(P)\eta(P^{-1}) \quad \text{and} \quad i[yx] = \sum \eta(P)\xi(P^{-1}).$$

Since the sums run over an entire group, these expressions are identical.

Notice that Lemma 2 implies the following cyclic property:

$$i[xyz] = i[zxy] = i[yzx],$$

for any elements x, y, z of the group algebra.

We now repeat the definition of the adjoint operation and prove two results.

DEFINITION. For any element $x = \sum \xi(P)P$ in $A(S_N)$, the *adjoint element* is defined to be

$$x^\dagger = \sum \xi^*(P)P^{-1},$$

where * denotes the complex conjugate.

LEMMA 3. *For any x and y in $A(S_N)$, $(xy)^\dagger = y^\dagger x^\dagger$.*

Proof. Defining x and y as before,

$$(xy)^\dagger = \left[\sum_P \sum_{P'} \xi(P)\eta(P')PP'\right]^\dagger = \sum_P \sum_{P'} \xi^*(P)\eta^*(P')P'^{-1}P^{-1}$$

$$= \left[\sum_{P'} \eta^*(P')P'^{-1}\right]\left[\sum_P \xi^*(P)P^{-1}\right] = y^\dagger x^\dagger.$$

LEMMA 4. *For any x in $A(S_N)$ other than the null, $i[xx^\dagger] > 0$.*

Proof. If $x = \sum \xi(P)P$, then $x^\dagger = \sum \xi^*(P)P^{-1}$, so that

$$i[xx^\dagger] = \sum_P |\xi(P)|^2 > 0,$$

if at least one coefficient $\xi(P)$ is nonzero.

We shall make frequent use of two properties of the tableau operators $R_r{}^D$, $C_r{}^D$, and $E_r{}^D$. These are proved in Rutherford (1948, p. 12, pp. 20 ff.), so they are quoted here without proof.

LEMMA 5. *For every D, r, and s,*

$$p^D_{rs} R_s{}^D = R_r{}^D p^D_{rs} \quad \textit{and} \quad p^D_{rs} C_s{}^D = C_r{}^D p^D_{rs},$$

so that

$$p^D_{rs} E_s{}^D = E_r{}^D p^D_{rs}.$$

LEMMA 6. *For every D, D', r, s, and every x in $A(S_N)$,*

$$E_r{}^D x E_s^{D'} = \delta^{DD'} E^D_{rs} \cdot \theta^D i[E^D_{sr} x],$$

where

$$\theta^D = [N!/d^D] > 0, \quad \textit{and} \quad E^D_{rs} = p^D_{rs} E_s{}^D = E_r{}^D p^D_{rs}.$$

B. Lemmas Concerning the Matric Bases

LEMMA 7. *If the numbers $N-1$ and N are on different rows and different columns in a standard tableau $T_r{}^D$ containing N numbers, then $E_r{}^D e_r^{D,2}$ does not contain the transposition $(N-1, N)$, i.e., the coefficient of $(N-1, N)$ in $E_r{}^D e_r^{D,2}$ is zero.*

Proof. The element $e_r^{D,2}$ does not operate on the numbers $N-1$ and N. Therefore, if $E_r{}^D e_r^{D,2}$ were to contain $(N-1, N)$, $E_r{}^D$ would have to contain a permutation of the form $(N-1, N)\tilde{p}$, where $\tilde{p}$ is a permutation which does not affect $N-1$ or N. We shall show that $E_r{}^D$ can contain no such permutation.

There are two forms possible for $T_r{}^D$, namely

$$\begin{array}{l} \quad\vdots \qquad \vdots \\ \ldots k \ldots N-1 \\ \quad\vdots \\ \ldots N \end{array} \tag{29}$$

and

$$\begin{array}{l} \quad\vdots \quad \vdots \\ \ldots k \ldots N \\ \quad\vdots \\ \ldots N-1 \end{array} \tag{30}$$

It is sufficient to consider only the former. With T_r^D of the form (29), E_r^D will contain only permutations of the form $\tilde{r}r_{N-1}r_N c_N c_{N-1}\tilde{c}$, where r_N is a row permutation for the row containing N, etc., and $\tilde{r}$, $\tilde{c}$ are permutations which do not operate on $N-1$ or N.

If

$$\tilde{r}r_{N-1}r_N c_N c_{N-1}\tilde{c} = (N-1, N)\tilde{p},$$

then

$$r_{N-1}r_N c_N c_{N-1} = (N-1, N)\tilde{r}^{-1}\tilde{p}\tilde{c}^{-1} = (N-1, N)\tilde{q}, \tag{31}$$

where $\tilde{q}$ does not operate on $N-1$ or N. We will prove that Eq. (31) is impossible.

According to Eq. (31), $r_{N-1}r_N c_N c_{N-1}$ must be a permutation in which N is replaced by $N-1$, and $N-1$ by N. We know from the form (29) of the tableau, however, that $r_{N-1}r_N c_N c_{N-1}$ has the form

$$(\ldots k \ldots N-1)(\ldots N)(\ldots k \ldots N)(\ldots N-1), \tag{32}$$

where the dots represent numbers other than k, $N-1$, or N. Now, because of the form of the tableau, *no two of these permutations can share any numbers other than* k, $N-1$, *and* N. Thus if, in $c_N = (\ldots k \ldots N)$, N is replaced by a number l other than k or $N-1$, the product $r_{N-1}r_N c_N c_{N-1}$ will be a permutation $(\ldots Nl)$ because neither r_{N-1} nor r_N will operate on l. Consequently, if any permutation of the form (32) can satisfy (31), it will be one in which the numbers represented by dots play no part at all. We may just as well consider the simpler tableau

$$T_r^D = \begin{array}{|c|c|} \hline k & N-1 \\ \hline N & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array}$$

But then (neglecting a numerical factor)

$$\begin{aligned} E_r^D &= [I + (k, N-1)][I - (k, N)] \\ &= I + (k, N-1) - (k, N) - (k, N, N-1). \end{aligned}$$

We have proved that E_r^D can contain no permutation of the form $(N-1, N)\tilde{p}$ if T_r^D is of the form (29). The proof for (30) is similar.

LEMMA 8.

$$i[E_r^D M_r^D e_r^{D,m}] = k \cdot i[E_r^{D,m} e_r^{D,m}], \qquad k > 0,$$

for every D *and* r, *and for* $m = 1$ *or* $m = 2$.

Proof. We deal here with operators defined in terms of a single standard tableau and its m-chain. We therefore drop the superscripts and subscripts, and denote E_r^D by E, M_r^D by M, $e_r^{D,m}$ by e^-, and $E_r^{D,m}$ by E^-.

For a 1-chain, it can be shown (Rutherford, 1948, p. 28) that

$$\begin{aligned} E &= E^- + (\text{terms operating on } N) \\ &= E^- + t_N. \end{aligned}$$

Therefore, $i[Ee^-] = i[E^-e^-] + i[t_Ne^-]$. The last term is zero because e^- does not operate on N, and t_N is made up *only* of permutations that operate on N. Thus t_Ne^- cannot contain the identity. This proves the theorem when $m = 1$.

For a 2-chain, there are three cases:

(1) If $N - 1$ and N appear on the same row of T_r^D, then $M_r^DR_r^D = R_r^D$ because R_r^D contains the idempotent M_r^D and is a group sum. Thus

$$\begin{aligned} i[EMe^-] &= i[RCMe^-] = i[RCe^-M] = i[MRCe^-] \\ &= i[RCe^-] = i[Ee^-]. \end{aligned}$$

We have used Lemma 2 and the fact that M commutes with e^-.

(2) If $N - 1$ and N appear on the same column of T_r^D, the argument is similar: $C_r^DM_r^D = C_r^D$, so that

$$i[EMe^-] = i[RCMe^-] = i[RCe^-] = i[Ee^-].$$

(3) If $N - 1$ and N occur on different rows and different columns in T_r^D, then

$$\begin{aligned} i[EMe^-] &= i[E \cdot \tfrac{1}{2}\{I \pm (N-1, N)\} \cdot e^-] \\ &= \tfrac{1}{2}i[Ee^-] \pm \tfrac{1}{2}i[E \cdot (N-1, N) \cdot e^-]. \end{aligned}$$

The last term contains $i[E \cdot (N-1, N) \cdot e^-] = i[Ee^- \cdot (N-1, N)]$, which is zero unless Ee^- contains $(N - 1, N)$. We proved in Lemma 7 that this is impossible.

In all three cases, $i[EMe^-] = k \cdot i[Ee^-]$, where $k > 0$. By an argument exactly parallel to that for 1-chains, it can be shown that $i[Ee^-] = i[E^-e^-]$. This proves the theorem for $m = 2$.

C. Existence Proofs

Our purpose in this section is to show that none of the matric basis elements vanish or blow up. The definitions involve factors ρ_r^D in the denominators. We begin by proving that these quantities are never zero. As a by-product, we are able to show that the diagonal elements of the matric basis are idempotent and Hermitian.

THEOREM 1. *For any D and r, and for $m = 1$ or $m = 2$:*
(a) $\rho_r{}^D \overset{\mathrm{d}}{=} i[C_r{}^D E_r{}^D M_r{}^D e_r^{D,m}] = i[C_r{}^D R_r{}^D C_r{}^D M_r{}^D e_r^{D,m}] > 0$;
(b) $i[E_r{}^D M_r{}^D e_r^{D,m}] \neq 0$;
(c) $e_r{}^D$ *is idempotent and self-adjoint.*

Proof. The proof is by induction. Using the notation of the previous lemma, we *assume* that

$$i[E^- e^-] \neq 0, \qquad e^- e^- = e^-, \qquad e^{-\dagger} = e^-,$$

then show that these properties recur: that

$$i[Ee] \neq 0, \qquad ee = e, \qquad e^\dagger = e, \tag{33}$$

and also that $i[CEMe^-] > 0$, $i[EMe^-] \neq 0$. This is shown in five steps.

(1) We assume that $i[E^- e^-] \neq 0$, so that $i[EMe^-] \neq 0$ by Lemma 8. This is the induction for part (b). Therefore, $x \overset{\mathrm{d}}{=} EMe^-$ is not the null, and $i[xx^\dagger] > 0$ by Lemma 4.

(2)

$$i[CEMe^-] = i[CRCMe^-] = i[CRRCMe^-]/o_{\mathfrak{R}}{}^D = i[RCMe^- CR]/o_{\mathfrak{R}}{}^D,$$

using Lemmas 1 and 2. By construction, M is idempotent and commutes with e^-: $Me^- = Me^- M$. In addition, we assume that e^- is idempotent, so $Me^- = Me^- e^- M$, and

$$i[CEMe^-] = i[(RCMe^-)(e^- MCR)]/o_{\mathfrak{R}}{}^D.$$

Furthermore, $(RCMe^-)^\dagger = e^{-\dagger} M^\dagger C^\dagger R^\dagger$. We assume that $e^{-\dagger} = e^-$, and $M^\dagger = M$, $C^\dagger = C$, $R^\dagger = R$ by construction. Thus

$$(RCMe^-)^\dagger = (e^- MCR)$$

and $i[CEMe^-] = i[(RCMe^-)(RCMe^-)^\dagger]/o_{\mathfrak{R}}{}^D = i[xx^\dagger]/o_{\mathfrak{R}}{}^D > 0$. This is the induction for part (a). We have yet to justify Eq. (33).

(3) Since $\rho = i[CEMe^-] \neq 0$ by (2), the quantity

$$i[Ee] = i[Ee^- MCEMe^-]/(\theta\rho)$$

is defined. But

$$\begin{aligned} Ee^- MCE &= E \cdot \theta i[Ee^- MC] && \text{by Lemma 6,} \\ &= E \cdot \theta i[CEMe^-] && \text{using Lemma 2,} \\ &= E \cdot \theta\rho, \end{aligned}$$

so that

$$i[Ee] = i[EMe^-] \neq 0 \qquad \text{by (1).}$$

(4) Assuming that e^- is idempotent,

$$\begin{aligned} ee &= e^-MCEMe^-e^-MCEMe^-/(\theta\rho)^2 \\ &= e^-MC\cdot EMe^-CE\cdot Me^-/(\theta\rho)^2 \\ &= e^-MC\cdot E\theta\rho\cdot Me^-/(\theta\rho)^2 \qquad \text{[as in (3)]} \\ &= e^-MCEMe^-/(\theta\rho) \\ &= e. \end{aligned}$$

(5) $e^\dagger = (e^-MCRCMe^-)^\dagger/(\theta\rho)$, since θ and ρ are real. We assume that $e^{-\dagger} = e^-$, so that

$$\begin{aligned} e^\dagger &= e^-M^\dagger C^\dagger R^\dagger C^\dagger M^\dagger e^-/(\theta\rho) \\ &= e^-MCRCMe^-/(\theta\rho) \\ &= e. \end{aligned}$$

This completes the induction scheme. We now prove that the induction has a base.

For a 1-chain, $E_r^{D,N-1} = e_r^{D,N-1} = I$, so that $e_r^{D,N-1}$ is idempotent and self-adjoint, and

$$i[E_r^{D,N-1}e_r^{D,N-1}] = i[I] = 1 \neq 0.$$

For 2-chains, $E_r^{D,N-2} = e_r^{D,N-2} = [I \pm (1,2)]/2$, so that $e_r^{D,N-2}$ is idempotent and self-adjoint, and

$$i[E_r^{D,N-2}e_r^{D,N-2}] = i[e_r^{D,N-2}] = \tfrac{1}{2} \neq 0. \qquad \text{Q.E.D.}$$

THEOREM 2. *None of the elements e_{rs}^D is the null.*

Proof. We prove that ${E_r}^D e_{rs}^D {E_s}^D$ does not vanish. This is

$${E_r}^D e_{rs}^D {E_s}^D = \underline{{E_r}^D e_r^{D,m} {M_r}^D {C_r}^D {E_r}^D} p_{rs}^D {M_s}^D e_s^{D,m} {E_s}^D/[\theta^D({\rho_r}^D{\rho_s}^D)^{1/2}].$$

The underlined part is ${E_r}^D\theta^D{\rho_r}^D$ by Lemmas 6 and 2 [the argument is similar to that in step (3) of Theorem 1], so that

$${E_r}^D e_{rs}^D {E_s}^D = {E_r}^D p_{rs}^D {M_s}^D e_s^{D,m} {E_s}^D({\rho_r}^D/{\rho_s}^D)^{1/2},$$

or

$$\begin{aligned} {E_r}^D e_{rs}^D {E_s}^D &= p_{rs}^D {E_s}^D {M_s}^D e_s^{D,m} {E_s}^D({\rho_r}^D/{\rho_s}^D)^{1/2} \\ &= p_{rs}^D {E_s}^D \theta^D i[{E_s}^D {M_s}^D e_s^{D,m}]({\rho_r}^D/{\rho_s}^D)^{1/2}. \end{aligned}$$

It was shown in Theorem 1 that $\rho_r{}^D$, $\rho_s{}^D$, and $i[E_s{}^D M_s{}^D e_s^{D,m}]$ are nonzero. Also $\theta^D > 0$ by Lemma 6 and $p_{rs}^D E_s{}^D$ is not the null.[3] This completes the proof.

We have now proved that the definitions (18)–(25) of the orthogonal matric bases yield existing, nonvanishing operators.

D. Multiplicative Properties

Theorem 1 has already shown that the diagonal elements $e_r{}^D = e_{rr}^D$ are idempotent. This fact, and the two lemmas that follow, are enough to establish the matric basis multiplication relation.

First we must show that

$$M_r{}^D e_r^{D,m} e_s^{D,m} M_s{}^D = \delta_{rs} M_r{}^D e_r^{D,m}. \tag{34}$$

It is clear to begin with that

$$M_r{}^D e_r^{D,m} e_r^{D,m} M_r{}^D = M_r{}^D e_r^{D,m},$$

because $M_r{}^D$ and $e_r^{D,m}$ are idempotent and commuting. It remains only to show that

$$M_r{}^D e_r^{D,m} e_s^{D,m} M_s{}^D = \text{null} \qquad \text{if} \quad r \neq s.$$

This is the purpose of Lemmas 9 and 10.

Lemma 9 (For 1-chains). *Suppose that two standard tableaux $T_r{}^D$ and $T_s{}^D$ belong to the same diagram D and differ in the position of the highest number N. Then $T_r^{D,1}$ and $T_s^{D,1}$ belong to different diagrams, and*

$$E_r^{D,1} x E_s^{D,1} = \text{null} = E_s^{D,1} x E_r^{D,1}$$

for every x in the group algebra.

(For 2-chains). *Let two standard tableaux $T_r{}^D$ and $T_s{}^D$ belonging to the same diagram D differ in the position of at least one of the two highest numbers $N-1$ and N. Then either $M_r{}^D M_s{}^D = \text{null} = M_s{}^D M_r{}^D$ or $T_r^{D,2}$ and $T_s^{D,2}$ belong to different diagrams. In the latter case,*

$$E_r^{D,2} x E_s^{D,2} = \text{null} = E_s^{D,2} x E_r^{D,2}$$

for every x in $A(S_N)$.

---

[3] From Rutherford (1948, p. 14), the coefficient of I in $E_s{}^D$ is 1. Hence this is also the coefficient of p_{rs}^D in E_{rs}^D. Since E_{rs}^D contains at least one permutation with a nonzero coefficient, it cannot be the null.

Proof. For 1-chains, it is obvious that $T_r^{D,1}$ and $T_s^{D,1}$ will belong to different diagrams. The conclusion follows from Lemma 6.

For 2-chains, the argument is similar except when $T_r{}^D = (N-1, N)T_s{}^D$. In such a case, $T_r^{D,2}$ and $T_s^{D,2}$ will belong to the same diagram, but we have defined $M_r{}^D$ and $M_s{}^D$ such that one will symmetrize $N-1$ and N, and the other will antisymmetrize them. In this case,

$$M_r{}^D M_s{}^D = \text{null} = M_s{}^D M_r{}^D.$$

LEMMA 10. *Let $T_r{}^D$ and $T_s{}^D$ be different standard tableaux belonging to the same diagram D. Then*

$$M_r{}^D e_r^{D,m} e_s^{D,m} M_s{}^D = \text{null}.$$

Proof. If $T_r{}^D$ and $T_s{}^D$ differ in the positions of their highest one (for 1-chains) or two (for 2-chains) numbers, then Lemma 9 applies directly, and since $M_s{}^D$ commutes with $e_t^{D,m}$ for every t,

$$\begin{aligned} M_r{}^D e_r^{D,m} e_s^{D,m} M_s{}^D &= M_r{}^D M_s{}^D e_r^{D,m} e_s^{D,m} \\ &= (\text{number})\underline{M_r{}^D M_s{}^D} e_r^{D,2m} M_r^{D,m} C_r^{D,m} \\ &\quad \times \underline{E_r^{D,m} M_r^{D,m} e_r^{D,2m} e_s^{D,2m} M_s^{D,m} C_s^{D,m} E_s^{D,m}} M_s^{D,m} e_s^{D,2m} \end{aligned}$$

where one or the other of the underlined factors is the null.

Otherwise, there is a number k such that removal of the highest km numbers from $T_r{}^D$ and $T_s{}^D$ results in tableaux $T_r^{D,km}$ and $T_s^{D,km}$ differing in the positions of their highest m numbers. Then recursive substitution gives

$$\begin{aligned} &M_r{}^D e_r^{D,m} e_s^{D,m} M_s{}^D \\ &\quad = (\text{number}) M_r{}^D e_r^{D,2m} M_r^{D,m} C_r^{D,m} \\ &\qquad \times E_r^{D,m} M_r^{D,m} e_r^{D,2m} e_s^{D,2m} M_s^{D,m} C_s^{D,m} E_s^{D,m} M_s^{D,m} e_s^{D,2m} M_s{}^D \\ &\quad = \cdots \\ &\quad = (\text{number}) M_r{}^D \cdots \underline{M_r^{D,km} e_r^{D,(k+1)m} e_s^{D,(k+1)m} M_s^{D,km}} \cdots M_s{}^D \end{aligned}$$

where the underlined factor is the null, by the argument given above. This proves the lemma.

Lemmas 6 and 10 and Theorem 1(c) put us in position to show how the elements e_{rs}^D multiply.

THEOREM 3. $e_{rs}^D e_{tu}^{D'} = \delta^{DD'} \delta_{st} e_{ru}^D$.

Proof.

$$\begin{aligned} e^D_{rs} e^{D'}_{tu} &= [\theta^D \theta^{D'} (\rho_r{}^D \rho_s{}^D \rho^{D'}_t \rho^{D'}_u)^{1/2}]^{-1} \\ &\quad \times e^{D,m}_r M_r{}^D p^D_{rs} C_s{}^D \underline{E_s{}^D M_s{}^D e^{D,m}_s e^{D',m}_t M^{D'}_t p^{D'}_{tu} C^{D'}_u E^{D'}_u} M^{D'}_u e^{D',m}_u \end{aligned}$$

where the underlined factor vanishes if $D \neq D'$, by Lemma 6. Therefore,

$$\begin{aligned} e^D_{rs} e^{D'}_{tu} &= \delta^{DD'} e^{D,m}_r M_r{}^D p^D_{rs} C_s{}^D E_s{}^D \underline{M_s{}^D e^{D,m}_s e^{D,m}_t M_t{}^D} p^D_{tu} C_u{}^D E_u{}^D M_u{}^D e^{D,m}_u \\ &\quad \times (\theta^D)^{-2} (\rho_r{}^D \rho_s{}^D \rho_t{}^D \rho_u{}^D)^{-1/2}. \end{aligned}$$

By Lemma 10, the underlined factor is $\delta_{st} M_s{}^D e^{D,m}_s$, so

$$\begin{aligned} e^D_{rs} e^{D'}_{tu} &= \delta^{DD'} \delta_{st} e^{D,m}_r M_r{}^D p^D_{rs} C_s{}^D E_s{}^D M_s{}^D e^{D,m}_s p^D_{su} C_u{}^D E_u{}^D M_u{}^D e^{D,m}_u \\ &\quad \times (\theta^D)^{-2} (\rho_s{}^D)^{-1} (\rho_r{}^D \rho_u{}^D)^{-1/2}. \end{aligned}$$

Using the fact that $p^D_{su} C_u{}^D E_u{}^D = C_s{}^D E_s{}^D p^D_{su}$, from Lemma 5,

$$\begin{aligned} e^D_{rs} e^{D'}_{tu} &= \delta^{DD'} \delta_{st} e^{D,m}_r M_r{}^D p^D_{rs} C_s{}^D \underline{E_s{}^D M_s{}^D e^{D,m}_s C_s{}^D E_s{}^D} p^D_{su} M_u{}^D e^{D,m}_u \\ &\quad \times (\theta^D)^{-2} (\rho_s{}^D)^{-1} (\rho_r{}^D \rho_u{}^D)^{-1/2}. \end{aligned}$$

By Lemmas 6 and 2, the underlined part is

$$E_s{}^D \theta^D i[E_s{}^D M_s{}^D e^{D,m}_s C_s{}^D] = E_s{}^D \theta^D i[C_s{}^D E_s{}^D M_s{}^D e^{D,m}_s] = E_s{}^D \theta^D \rho_s{}^D,$$

so that

$$\begin{aligned} e^D_{rs} e^{D'}_{tu} &= \delta^{DD'} \delta_{st} e^{D,m}_r M_r{}^D p^D_{rs} C_s{}^D E_s{}^D p^D_{su} M_u{}^D e^{D,m}_u / [\theta^D (\rho_r{}^D \rho_u{}^D)^{1/2}] \\ &= \delta^{DD'} \delta_{st} e^{D,m}_r M_r{}^D p^D_{rs} p^D_{su} C_u{}^D E_u{}^D M_u{}^D e^{D,m}_u / [\theta^D (\rho_r{}^D \rho_u{}^D)^{1/2}] \\ &= \delta^{DD'} \delta_{st} e^{D,m}_r M_r{}^D p^D_{ru} C_u{}^D E_u{}^D M_u{}^D e^{D,m}_u / [\theta^D (\rho_r{}^D \rho_u{}^D)^{1/2}] \\ &= \delta^{DD'} \delta_{st} e^D_{ru}. \end{aligned}$$

This proves the theorem.

E. Orthogonal Operator Bases for Every Irreducible Representation

It follows from Theorem 3 that a matric basis $\{e^D_{rs} \mid \text{all } D, r, s\}$ consists of $\sum_D (d^D)^2 = N!$ linearly independent elements. Thus the YK- and Serber-adapted matric bases introduced here *span the entire group algebra.*

Furthermore, they have been defined in such a way that

$$\begin{aligned} e^{D\dagger}_{rs} &= [(E_r{}^D M_r{}^D e^{D,m}_r)^\dagger p^D_{rs} (E_s{}^D M_s{}^D e^{D,m}_s)]^\dagger / (k_r{}^D k_s{}^D)^{1/2} \\ &= [(E^D_s M_s{}^D e^{D,m}_s)^\dagger p^D_{sr} (E_r{}^D M_r{}^D e^{D,m}_r)] / (k_s{}^D k_r{}^D)^{1/2} \\ &= e^D_{sr}. \end{aligned}$$

Thus these matric bases have the adjoint property (7). It follows that a subset

$$B_s{}^D = \{e^D_{rs} \mid s \text{ fixed, all } r\}$$

spans a carrier space for an *orthogonal* representation of S_N. We say that $B_s{}^D$ is an operator basis for an orthogonal representation, or for short, an *orthogonal operator basis.*

Now $B_s{}^D$ consists of elements $e^D_{rs} = e^D_{rs} e^D_{ss}$, with s fixed, spanning a left ideal generated by the idempotent e^D_{ss}. As a matter of fact, this left ideal is *minimal,* for we now show that e^D_{ss} is a *primitive idempotent.*

THEOREM 4. *For any D, D', r, and s, and any element x in the group algebra,*

$$e_r{}^D x e^{D'}_s = e^D_{rr} x e^{D'}_{ss} = \delta^{DD'} \lambda(x) e^D_{rs},$$

where $\lambda(x)$ is a number that depends on x.

Proof.

$$\begin{aligned} e_r{}^D x e^{D'}_s &= e^{D,m}_r M_r{}^D C_r{}^D \underline{E_r{}^D M_r{}^D e^{D,m}_r x e^{D',m}_s M^{D'}_s C^{D'}_s E^{D'}_s} M^{D'}_s e^{D',m}_s \\ &\quad \times (\theta^D \theta^{D'} \rho_r{}^D \rho^{D'}_s)^{-1}. \end{aligned}$$

Applying Lemma 6 to the underlined portion,

$$\begin{aligned} e_r{}^D x e^{D'}_s &= \delta^{DD'} e^{D,m}_r M_r{}^D C_r{}^D p^D_{rs} E_s{}^D M_s{}^D e^{D,m}_s \\ &\quad \times i[p^D_{sr} E_r{}^D M_r{}^D e^{D,m}_r x e^{D,m}_s M_s{}^D C_s{}^D](\theta^D \rho_r{}^D \rho_s{}^D)^{-1} \\ &= \delta^{DD'} \cdot (\text{number}) \cdot e^D_{rs}. \end{aligned}$$

We have as a special case of this result

$$e_r{}^D x e_r{}^D = \lambda(x) e_r{}^D$$

for arbitrary x. Thus $e_r{}^D$ has the property (14): the diagonal elements of the matric bases are *primitive idempotents.*

These idempotents, unlike the Young idempotents, generate the minimal left ideals occurring in the decomposition of the group algebra. This we prove by showing that the identity, the generating unit of the whole group algebra, decomposes as the sum of the linearly independent elements $e_r{}^D$, which generate minimal left ideals.

THEOREM 5. $\sum_D \sum_r e_r{}^D = \sum_D \sum_r e^D_{rr} = I.$

Proof. Let $T = \sum_D \sum_r e_{rr}^D$. Since the matric basis spans $A(S_N)$, an arbitrary element x can be expanded in the form

$$x = \sum_D \sum_r \sum_s \xi_{rs}^D e_{rs}^D .$$

It follows that

$$\begin{aligned} xT &= \sum_D \sum_r \sum_s \xi_{rs}^D e_{rs}^D \sum_{D'} \sum_t e_{tt}^{D'} = \sum_D \sum_r \sum_s \sum_{D'} \sum_t \xi_{rs}^D (e_{rs}^D e_{tt}^{D'}) \\ &= \sum_D \sum_r \sum_s \sum_{D'} \sum_t \xi_{rs}^D (\delta^{DD'} \delta_{st} e_{rs}^D) = \sum_D \sum_r \sum_s \xi_{rs}^D e_{rs}^D \\ &= x, \end{aligned}$$

for arbitrary x. Similarly, $Tx = x$. It follows that $T = I$.

It should be noted that this theorem cannot be proved with Young idempotents $E_r{}^D$ in place of the $e_r{}^D$. This is because $E_r{}^D E_s{}^D \neq \delta_{rs} E_r{}^D$, in general.

VI. Generating Antisymmetric Many-Electron Wavefunctions from Conjugate Young Diagrams

A. Construction by Means of Wigner Operators

Once a set of pure spin eigenfunctions θ_i has been generated, corresponding antisymmetric space–spin functions can be readily constructed. Since $\mathscr{S}^2$ and $\mathscr{S}_z$ commute with the space–spin antisymmetrizer

$$\mathscr{A} = (N!)^{-1} \sum_P \varepsilon(P) P = (N!)^{-1} \sum_P (-1)^P P,$$

it is apparent that any wavefunction

$$\Psi_i = \mathscr{A}\{f(\text{space})\theta_i(\text{spin})\}, \tag{35}$$

if it is nonzero, is not only antisymmetric, but is also an eigenfunction of $\mathscr{S}^2$ and $\mathscr{S}_z$. Here the primitive space function f is arbitrary. If it is a product of orbitals, Ψ_i decomposes into a superposition of Slater determinants when θ_i is expressed in terms of products of α's and β's.

The antisymmetrizer in Eq. (35) masks the relationship between the space and spin components in Ψ_i. Suppose that there are $d(N, S)$ spin eigenfunctions $\theta_i(N, S, M)$ for given N, S, and M. Then these functions span an irreducible matrix representation $[P]^{NS}$ of S_N: for any permutation P transforming the spin coordinates of electron 1, 2, ..., N, one has

$$P\theta_i(N, S, M) = \sum_j \theta_j(N, S, M)[P]_{ji}^{NS}. \tag{36}$$

This will be called the *spin representation* of S_N. Using this relation in Eq. (35), we can rewrite Ψ_i in the form (Kotani *et al.*, 1955, Ch. 1, Sects. 5 and 6)

$$\Psi_i(N, S, M) = [d(N, S)]^{-1} \sum_j f_j(N, S; i)\theta_j(N, S, M), \qquad (37)$$

where

$$f_j(N, S; i) = [d(N, S)/N!] \sum_P \varepsilon(P)[P]^{NS}_{ji}(Pf). \qquad (38)$$

Equation (37) shows that Ψ_i is a sum of terms, each of which is the product of a spin eigenfunction and some kind of space function. The space function, shown in Eq. (38), is generated from the primitive function f by a Wigner operator. As a consequence, such space functions form a basis for an irreducible representation of S_N, called the *space representation*: if $k = d(N, S)/N!$,

$$\begin{aligned} Pf_j(N, S; i) &= k \sum_{P'} \varepsilon(P')[P']^{NS}_{ji}(PP'f) \\ &= k \sum_{P''} \varepsilon(P^{-1}P'')[P^{-1}P'']^{NS}_{ji}(P''f) \\ &= k \sum_{P''} \varepsilon(P^{-1})\varepsilon(P'') \sum_m [P^{-1}]^{NS}_{jm}[P'']^{NS}_{mi}(P''f) \\ &= \varepsilon(P^{-1}) \sum_m \left\{ k \sum_{P''} \varepsilon(P'')[P'']^{NS}_{mi}(P''f) \right\} [P^{-1}]^{NS}_{jm} \\ &= \varepsilon(P) \sum_m f_m(N, S; i)[P^{-1}]^{NS}_{jm}. \end{aligned}$$

Comparison with Eq. (36) shows that when the spin functions transform under P according to the matrix $[P]^{NS}$, the space functions $f_j(N, S; i)$ transform according to the representation matrix $\widehat{[P]}^{NS}$, the elements of which are given by

$$\widehat{[P]}^{NS}_{ij} = \varepsilon(P)[P^{-1}]^{NS}_{ji},$$

which for orthogonal representations like those considered here, simplifies to

$$\widehat{[P]}^{NS}_{ij} = \varepsilon(P)[P]^{NS}_{ij}.$$

The spin and space irreps are reciprocal in such a way that Ψ_i is antisymmetric: they are said to be *dual.*[4]

---

[4] See Kotani *et al.* (1955, Ch. 1, Sects. 5 and 6). Wigner uses the term "associated" (see Wigner, 1959, p. 127, 258).

B. Construction by Means of Genealogical Matric Bases

When Ψ_i is constructed in the form (37), there is *no sum over $N!$ permutations* as there is in Eq. (35). Thus the construction from dual space and spin functions may be more convenient, *if these can be generated easily.* The construction of dual functions by means of Wigner operators has been discussed by Kotani *et al.* (1955), Harris (1967), and Sullivan (1968). Goddard (1967, 1968) has made extensive use of Young's orthogonal units for this purpose. But these methods of construction are not practical for large numbers of electrons, since *the generating operators involve sums over all permutations in S_N*.

We now discuss how dual spin and space functions are related in terms of Young diagrams and the genealogical matric bases defined earlier. We have mentioned that spin representations are labeled by diagrams with one or two rows. It turns out that space functions transform according to irreducible representations associated with diagrams having one or two *columns.*

A diagram obtained from another diagram by interchanging rows and columns is said to be *conjugate* to it. For example,

$$\begin{array}{|c|c|}\hline & \\ \hline & \\ \hline \\ \cline{1-1}\end{array} \quad \text{is conjugate to} \quad \begin{array}{|c|c|c|}\hline & & \\ \hline & \\ \cline{1-2}\end{array}$$

while

$$\begin{array}{|c|c|}\hline & \\ \hline \\ \cline{1-1}\end{array} \quad \text{is self-conjugate.}$$

Space and spin diagrams are conjugate. The fact that dual irreps of S_N belong to conjugate diagrams was first mentioned by Weyl (1950, p. 369), who gave a proof by tensor methods.

It is useful to define *conjugate standard tableaux* in a similar way. Let T_r^D be a standard tableau belonging to the diagram D. Then we shall denote by $T_{\tilde{r}}^{\tilde{D}}$ the standard tableau conjugate to T_r^D and belonging to the diagram $\tilde{D}$ conjugate to D; i.e., $T_{\tilde{r}}^{\tilde{D}}$ is obtained from T_r^D by changing rows into columns, and *vice versa.* For example,

$$\text{if} \quad T_r^D = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1}\end{array}, \qquad \text{then} \quad T_{\tilde{r}}^{\tilde{D}} = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 \\ \cline{1-1} 3 \\ \cline{1-1}\end{array}$$

Now suppose that θ is a pure-spin eigenfunction of $\mathscr{S}_z$ with eigenvalue M. Let D be an N-spin diagram corresponding to $\mathscr{S}^2$ eigenvalue $S \geq |M|$, and let e_{rt}^{D} be a matric basis element for diagram D of the type introduced in Section IV. Then the functions

$$\{\theta_r{}^D(t) = e_{rt}^{D}\theta \,|\, t \text{ fixed and arbitrary; } r = 1, 2, \ldots, d^D\}$$

span a spin irrep, Γ^D, of S_N. By construction, they are spin eigenfunctions with eigenvalues S and M. The function $\theta_r{}^D$ "belongs to the standard tableau $T_r{}^D$."

According to Weyl, the space representation dual to Γ^D belongs to the diagram $\tilde{D}$: call it $\Gamma^{\tilde{D}}$. Now the operator basis

$$\{e_{uv}^{\tilde{D}} \,|\, v \text{ fixed and arbitrary; } u = 1, 2, \ldots, d^{\tilde{D}}\}$$

spans an irrep belonging to diagram $\tilde{D}$, and thus equivalent to $\Gamma^{\tilde{D}}$. Therefore it must be possible to find a space primitive function f such that the functions

$$\{f_u{}^{\tilde{D}}(v) = e_{uv}^{\tilde{D}} f \,|\, v \text{ fixed and arbitrary; } u = 1, 2, \ldots, d^{\tilde{D}}\},$$

which "belong to the standard tableaux conjugate to the spin tableaux," span an irrep dual to the spin representation Γ^D.

It follows that there is an antisymmetric space–spin function of the form

$$\Psi(N, S, M) = \sum_u \sum_r \beta_{ur} f_u{}^{\tilde{D}}(v)\theta_r{}^D(t), \tag{39}$$

where the β_{ur} are numbers that may depend on D, v, and t. Clearly, Ψ will be an eigenfunction of $\mathscr{S}^2$ and $\mathscr{S}_z$, with eigenvalue S determined by the diagram D and eigenvalue M determined by the spin primitive θ.

We shall prove in the following pages the validity of Weyl's assertion when applied to irreps generated by genealogical matric bases; i.e., we shall prove that antisymmetric wavefunctions take the form (39) when expressed in terms of dual spin and space functions.

However, the special properties of the genealogical matric bases allow a simplification. We will show that a space–spin function of the form (39) is antisymmetric if and only if it reduces to a "conjugate–diagonal" sum, in which *dual space and spin functions are generated by matric basis elements belonging to conjugate standard tableaux*. That is, we will prove that (39) reduces to the form

$$\Psi = \sum_r \beta_{\tilde{r}r} f_{\tilde{r}}{}^{\tilde{D}}(v)\theta_r{}^D(t), \tag{40}$$

where v and t are arbitrary.

Thus the genealogical matric bases of Section IV are defined in such a way that the concepts of duality and conjugation are closely linked, resulting in especially simple expressions for antisymmetric wavefunctions. *Equation* (40) *represents the construction of antisymmetric spin states by means of space and spin operators that are read at sight from conjugate standard tableaux.*

C. Duality and the Genealogical Matric Bases: Proofs

Goddard [1967, p. 74, Eqs. (8)–(11)] showed that a wavefunction of the form (40) is antisymmetric when the generating operators are Young orthogonal units. Here we wish to show, using operator multiplication properties, that an antisymmetric wavefunction *necessarily* has the form (40) when written in terms of the two genealogical matric bases of Section IV.

So far, we have dealt only with matric bases having idempotents defined by the recursion

$$\begin{aligned} e_r{}^D &= (E_r{}^D M_r{}^D e_r^{D,m})^\dagger (E_r{}^D M_r{}^D e_r^{D,m})/k_r{}^D \\ &= e_r^{D,m} M_r{}^D C_r{}^D R_r{}^D C_r{}^D M_r{}^D e_r^{D,m}/(\theta^D \rho_r{}^D), \end{aligned}$$

where $k_r{}^D = \theta^D o_{\mathscr{R}}{}^D \rho_r{}^D$, in which $\rho_r{}^D = i[C_r{}^D E_r{}^D M_r{}^D e_r^{D,m}]$ and $o_{\mathscr{R}}{}^D =$ the order of the row group for any tableau belonging to diagram D.

There is another construction possible, however; namely, the recursion

$$\begin{aligned} \hat{e}_r{}^D &= (\hat{e}_r^{D,m} M_r{}^D E_r{}^D)(\hat{e}_r^{D,m} M_r{}^D E_r{}^D)^\dagger / \hat{k}_r{}^D \\ &= \hat{e}_r^{D,m} M_r{}^D R_r{}^D C_r{}^D R_r{}^D M_r{}^D \hat{e}_r^{D,m}/(\theta^D \hat{\rho}_r{}^D), \end{aligned} \tag{41}$$

where $\hat{k}_r{}^D = \theta^D o_{\mathscr{C}}{}^D \hat{\rho}_r{}^D$, in which $\hat{\rho}_r{}^D = i[E_r{}^D R_r{}^D M_r{}^D \hat{e}_r^{D,m}]$ and $o_{\mathscr{C}}{}^D =$ the order of the column group for any tableau belonging to diagram D.

It is easy to convince oneself that proofs exactly parallel to those of Section V apply to a matric basis $\{\hat{e}_{rs}^D\}$ defined in terms of the $\hat{e}_r{}^D$ and thus that $\hat{e}_{rs}^D$ is, in some sense, a *doppelgänger* of e_{rs}^D.[5] We will therefore take it for granted that the diagonal elements $\hat{e}_r{}^D$ exist and are idempotent; we shall refer to them as idempotents of the "mirror" matric basis.

---

[5] The relation between the idempotents $e_r{}^D$ and $\hat{e}_r{}^D$ is analogous to the relation between the two definitions of the Young idempotent $E_r{}^D$ seen in the literature: i.e., $E_r{}^D = R_r{}^D C_r{}^D$ in Boerner (1963) and Rutherford (1948); $E_r{}^D = C_r{}^D R_r{}^D$ in Weyl (1931) and Matsen (1964). It is even more closely analogous to the relation between the "structure projectors" $R_r{}^D C_r{}^D R_r{}^D$ and $C_r{}^D R_r{}^D C_r{}^D$ (see Matsen, 1964; Gallup, 1968, 1969).

As we shall see, the relation between $e_r{}^D$ and $\hat{e}_{\tilde{r}}{}^D$ is the key to proofs concerning duality in terms of the matric bases. We begin by proving an annihilation property.

LEMMA 11. *$e_r{}^D\hat{e}_s^{D'} = \hat{e}_s^{D'}e_r{}^D = \text{null}$ if $T_r{}^D \neq T_s^{D'}$.*

Proof.

$$\begin{aligned} &e_r{}^D\hat{e}_s^{D'} \\ &\quad = e_r^{D,m}M_r{}^DC_r{}^D\underline{R_r{}^DC_r{}^DM_r{}^De_r^{D,m}\hat{e}_s^{D',m}M_s^{D'}R_s^{D'}C_s^{D'}}R_s^{D'}M_s^{D'}\hat{e}_s^{D',m}/[\theta^D\theta^{D'}\rho_r{}^D\hat{\rho}_s^{D'}]. \end{aligned}$$

The underlined portion is the null unless $D' = D$ (Lemma 6). Thus

$$\begin{aligned} e_r{}^D\hat{e}_s^{D'} = \delta^{DD'}e_r^{D,m}M_r{}^DC_r{}^DR_r{}^DC_r{}^D&\underline{M_r{}^De_r^{D,m}\hat{e}_s^{D,m}M_s{}^D}R_s{}^DC_s{}^DR_s{}^DM_s{}^D\hat{e}_s^{D,m} \\ &\times [(\theta^D)^2\rho_r{}^D\hat{\rho}_s{}^D]^{-1}. \end{aligned}$$

But if $T_r{}^D$ and $T_s{}^D$ are different standard tableaux belonging to the same diagram,

$$M_r{}^De_r^{D,m}\hat{e}_s^{D,m}M_s{}^D = \text{null}.$$

The proof of this statement is exactly parallel to that used in Lemmas 9 and 10 (similarly for $\hat{e}_s^{D'}e_r{}^D$).

Equation (40) will now be verified by means of a lemma and two theorems. We prove that: (1) the direct product of two irreps of S_N contains the antisymmetric representation if and only if the two irreps belong to conjugate diagrams [Theorem 6], and (2) if bases for two irreps are generated by means of genealogical matric basis operators defined in terms of conjugate Young diagrams, a direct-product function is antisymmetric if and only if it has the form (40) [Theorem 7].

LEMMA 12. *Let Γ^D be an irreducible representation of S_N corresponding to the Young diagram D and spanned by the basis*

$$\{h_r{}^D = e_{rs}^D h \,|\, r = 1, 2, \ldots, d^D\},$$

where s is arbitrary and h is a suitable primitive function. Let $g^A = \mathscr{A}g$ be an antisymmetric function projected from a primitive function g, and let Γ^A denote the antisymmetric representation. Define $T_{\tilde{r}}{}^D$ to be the standard tableau conjugate to $T_r{}^D$. Then:

(a) *the direct product $\Gamma^A \otimes \Gamma^D$ is an irreducible representation belonging to the Young diagram $\tilde{D}$ conjugate to D;*

(b) *there exists a primitive function* $f = f(t)$ *such that* $\Gamma^{\tilde{D}} = \Gamma^{A} \otimes \Gamma^{D}$ *is spanned by the basis* $\{f_u{}^{\tilde{D}} = e_{ut}^{\tilde{D}} f \mid t$ *fixed and arbitrary;* $u = 1, 2, \ldots, d^{\tilde{D}}\}$, *and this basis has the property* $f_r^{\tilde{D}} = \alpha_r g^A h_r{}^D$, *where* α_r *is a nonzero number.*

Proof. By definition, $\Gamma^A \otimes \Gamma^D$ is spanned by the direct-product basis

$$\{g^A h_r{}^D \mid r = 1, 2, \ldots, d^D\}. \tag{42}$$

But Γ^A is one-dimensional, so $\Gamma^A \otimes \Gamma^D$ is irreducible and therefore labeled by some Young diagram D':

$$\Gamma^{D'} = \Gamma^A \otimes \Gamma^D.$$

This means that $\Gamma^{D'}$ is equivalent to the irrep spanned by the operator basis $\{e_{ut}^{D'} \mid t$ fixed and arbitrary; $u = 1, 2, \ldots, d^{D'}\}$. Consequently, there exists a primitive function f (depending on t) such that the functions

$$\{f_u^{D'} = e_{ut}^{D'} f \mid u = 1, 2, \ldots, d^{D'}\}$$

span $\Gamma^{D'}$. Comparing with (42), it follows that there are numbers μ_{ur} such that

$$f_u^{D'} = g^A \sum_r \mu_{ur} h_r{}^D. \tag{43}$$

By definition,

$$f_u^{D'} = e_{ut}^{D'} f = e_u^{D'} e_{ut}^{D'} f = e_u^{D'} f_u^{D'} \neq 0. \tag{44}$$

Thus, from Eq. (43),

$$e_u^{D'} g^A \sum_r \mu_{ur} h_r{}^D \neq 0 \tag{45}$$

for any choice of u. This equation is the key to the evaluation of the coefficients μ_{ur} occurring in Eq. (43).

It is possible to evaluate $e_u^{D'} g^A h_r{}^D$ directly. Consider the action of any group algebra element on $g^A h_r{}^D$: if $X = \sum_P \xi(P)P$, where $\xi(P)$ is a number depending on X and P, then

$$\begin{aligned} X g^A h_r{}^D &= \sum_P \xi(P) P(g^A h_r{}^D) = \sum_P \xi(P)(P g^A)(P h_r{}^D) \\ &= g^A \sum_P \xi(P)\varepsilon(P) P h_r{}^D, \end{aligned}$$

where $\varepsilon(P) = +1$ when P is even, -1 when P is odd. Thus, if $\mathscr{A}$ and $\mathscr{S}$ denote an antisymmetrizer and symmetrizer on given numbers,

$$\mathscr{A} g^A h_r{}^D = g^A \mathscr{S} h_r{}^D \quad \text{and} \quad \mathscr{S} g^A h_r{}^D = g^A \mathscr{A} h_r{}^D.$$

Since $e_u^{D'}$ is a product of various symmetrizers and antisymmetrizers,

$$e_u^{D'} g^A h_r{}^D = g^A \tilde{e}_u^{D'} h_r{}^D,$$

where $\tilde{e}_u^{D'}$ is obtained from $e_u^{D'}$ by replacing each symmetrizer by the corresponding antisymmetrizer and *vice versa*. This means that

$$e_u^{D'} g^A h_r{}^D = \phi_{\tilde{u}}^{\tilde{D}'} \cdot g^A \hat{e}_{\tilde{u}}^{\tilde{D}'} h_r{}^D, \tag{46}$$

where $\hat{e}_{\tilde{u}}^{\tilde{D}'}$ is a "mirror matric basis" idempotent as defined in Eq. (41), and

$$\phi_{\tilde{u}}^{\tilde{D}'} = \prod_{j=1}^{(N-m)/m} (\hat{k}_{\tilde{u}}^{\tilde{D}', jm} / k_u^{D', jm})$$

is finite and nonzero.

But $\hat{e}_{\tilde{u}}^{\tilde{D}'} h_r{}^D = \hat{e}_{\tilde{u}}^{\tilde{D}'} e_{rs}^D h = \hat{e}_{\tilde{u}}^{\tilde{D}'} e_r{}^D e_{rs}^D h = \hat{e}_{\tilde{u}}^{\tilde{D}'} e_r{}^D h_r{}^D = 0$ unless $T_{\tilde{u}}^{\tilde{D}'} = T_r{}^D$ (Lemma 11). Thus

$$e_u^{D'} g^A h_r{}^D = 0, \quad \text{unless } T_u^{D'} \text{ is conjugate to } T_r{}^D. \tag{47}$$

Comparing this result with Eq. (45), it is clear that D' must be the Young diagram conjugate to D. This proves part (a).

Combining Eqs. (43), (44), (46), and (47), it follows immediately that

$$f_{\tilde{u}}{}^{\tilde{D}} = e_{\tilde{u}}{}^{\tilde{D}} f_{\tilde{u}}{}^{\tilde{D}} = (\mu_{\tilde{u}u} \phi_u{}^D) \cdot g^A \hat{e}_u{}^D h_u{}^D \neq 0, \tag{48}$$

so that $\mu_{\tilde{u}u}$ is nonzero. We complete the proof by showing that

$$\hat{e}_u{}^D h_u{}^D = (\text{number}) \cdot h_u{}^D.$$

From Eqs. (48) and (43),

$$g^A \hat{e}_u{}^D h_u{}^D = g^A (\phi_u{}^D)^{-1} \sum_r (\mu_{\tilde{u}r} / \mu_{\tilde{u}u}) h_r{}^D$$

for *arbitrary* g^A. Therefore

$$\hat{e}_u{}^D h_u{}^D = (\phi_u{}^D)^{-1} \sum_r (\mu_{\tilde{u}r} / \mu_{\tilde{u}u}) h_r{}^D,$$

or, since $e_r{}^D h_r{}^D = h_r{}^D$,

$$\hat{e}_u{}^D h_u{}^D = (\phi_u{}^D)^{-1} \sum_r (\mu_{\tilde{u}r} / \mu_{\tilde{u}u}) e_r{}^D h_r{}^D. \tag{49}$$

Multiplying from the left by $e_p{}^D$,

$$e_p{}^D \hat{e}_u{}^D h_u{}^D = (\phi_u{}^D)^{-1} (\mu_{\tilde{u}p} / \mu_{\tilde{u}u}) e_p{}^D h_p{}^D = (\phi_u{}^D)^{-1} (\mu_{\tilde{u}p} / \mu_{\tilde{u}u}) h_p{}^D.$$

The left side of this equation is zero when $p \neq u$ (Lemma 11). On the right, $h_p{}^D$ is nonvanishing by assumption, and $\mu_{\tilde{u}u}$ and $(\phi_u{}^D)^{-1}$ have been shown to be nonzero. Therefore $\mu_{\tilde{u}p} = 0$ when $p \neq u$ and Eq. (49) yields

$$\hat{e}_u{}^D h_u{}^D = (\phi_u{}^D)^{-1} h_u{}^D.$$

Combining this result with Eq. (48),

$$f_{\tilde{u}}{}^{\tilde{D}} = \mu_{\tilde{u}u} g^A h_u{}^D.$$

This proves the lemma.

THEOREM 6. *Let Γ^D and $\Gamma^{D'}$ be real irreducible representations of S_N belonging to Young diagrams D and D′. Let Γ^A be the antisymmetric representation corresponding to the diagram $\{1^N\}$.*

Then the direct product $\Gamma^D \otimes \Gamma^{D'}$ contains Γ^A only if D and D′ are conjugate diagrams. If D and D′ are conjugate, $\Gamma^D \otimes \Gamma^{D'}$ contains Γ^A once only.

Proof. The number of times that Γ^A is contained in $\Gamma^D \otimes \Gamma^{D'}$ is

$$a(A, D \otimes D') = (N!)^{-1} \sum_P \chi^A(P)\chi^D(P)\chi^{D'}(P),$$

where $\chi^D(P)$ is the character of the permutation P in Γ^D.

However, it was shown in Lemma 12 that $\Gamma^A \otimes \Gamma^D$ is an irrep $\Gamma^{\tilde{D}}$ belonging to the diagram $\tilde{D}$ conjugate to D. Thus, $\chi^A(P)\chi^D(P) = \chi^{\tilde{D}}(P)$. Using the orthogonality property of real simple characters,

$$a(A, D \otimes D') = (N!)^{-1} \sum_P \chi^{\tilde{D}}(P)\chi^{D'}(P) = \delta^{\tilde{D}D'}.$$

THEOREM 7. *Let f and θ be space and spin primitive functions, respectively and define*

$$\{f_r{}^D = e_{rt}^D f \,|\, t \textit{ fixed and arbitrary};\ r = 1, 2, \ldots, d^D\},$$
$$\{\theta_r{}^D = e_{ru}^D \theta \,|\, u \textit{ fixed and arbitrary};\ r = 1, 2, \ldots, d^D\},$$

in terms of the matric bases $\{e_{rs}^D\}$ defined previously. Define $T_{\tilde{r}}^{\tilde{D}}$ to be the standard tableau conjugate to $T_r{}^D$.

Then the direct-product space–spin function

$$\Psi = \sum_r \sum_s \beta_{rs} f_{\tilde{r}}{}^{\tilde{D}} \theta_s{}^D$$

is antisymmetric if and only if $\beta_{rs} = \delta_{rs} \beta_{rr}$.

Proof. It is clear from Lemma 12 and its proof that there exist space primitives g and h such that

$$f_{\tilde r}^{\bar D} = \alpha_r g^A h_r{}^D,$$

where α_r is nonzero, $g^A = \mathscr{A} g$, and $h_r{}^D = e_{rv}^D h$, for some choice of v. Thus

$$\Psi = g^A \sum_r \sum_s \gamma_{rs} h_r{}^D \theta_s{}^D, \tag{50}$$

where $\gamma_{rs} = \alpha_r \beta_{rs}$.

It follows that

$$\begin{aligned} \mathscr{A}\Psi &= (N!)^{-1} \sum_P \varepsilon(P)(Pg^A) \sum_r \sum_s \gamma_{rs}(Ph_r{}^D)(P\theta_s{}^D) \\ &= (N!)^{-1} g^A \sum_P \sum_r \sum_s \gamma_{rs} \sum_i h_i{}^D [P]_{ir}^D \sum_j \theta_j{}^D [P]_{js}^D \\ &= g^A \sum_i \sum_j h_i{}^D \theta_j{}^D \cdot (N!)^{-1} \sum_r \sum_s \gamma_{rs} \sum_P [P]_{ir}^D [P]_{js}^D . \end{aligned}$$

Since the representation is orthogonal, $\sum_P [P]_{ir}^D [P]_{js}^D = (N!/d^D)\delta_{ij}\,\delta_{rs}$, and

$$\mathscr{A}\Psi = g^A \sum_i h_i{}^D \theta_i{}^D \cdot (d^D)^{-1} \sum_r \gamma_{rr}. \tag{51}$$

If Ψ is antisymmetric, it must be that $\mathscr{A}\Psi$ is proportional to Ψ. Comparing Eqs. (50) and (51) and considering the linear independence of the space–spin products $\{h_r{}^D \theta_s{}^D\}$, we see that this occurs only when $\gamma_{rs} = \delta_{rs}$ and Ψ is of the form $\Psi = g^A \sum_r h_r{}^D \theta_r{}^D$ or $\Psi = \sum_r (\alpha_r)^{-1} f_{\tilde r}^{\bar D} \theta_r{}^D$.

This completes the proof for the validity of Eq. (40).

ACKNOWLEDGMENTS

The author is indebted to Professor Klaus Ruedenberg, who suggested the thesis problem from which this research grew, and to Professor R. D. Poshusta for many helpful discussions concerning group algebra theory. Both provided considerable help and stimulation during the writing of the paper.

REFERENCES

BOERNER, H. (1963). "Representations of Groups." North-Holland Publ., Amsterdam.
COLEMAN, A. J. (1968). *Advan. Quantum Chem.* **4**, 83.
DIRAC, P. A. M. (1929). *Proc. Roy. Soc., Ser. A* **123**, 714.
ELLIOTT, J. P., HOPE, J., and JAHN, H. A. (1953). *Phil. Trans. Roy. Soc. London, Ser. A* **246**, 241.
GALLUP, G. A. (1968). *J. Chem. Phys.* **48**, 1752.

GALLUP, G. A. (1969). *J. Chem. Phys.* **50**, 1206.

GODDARD, W. A. (1967). *Phys. Rev.* **157**, 73, 81.

GODDARD, W. A. (1968). *J. Chem. Phys.* **48**, 450.

HARRIS, F. E. (1967). *Advan. Quantum Chem.* **3**, 61.

HEITLER, W., and LONDON, F. (1927). *Z. Phys.* **44**, 455.

JAHN, H. A. (1954). *Phys. Rev.* **96**, 989.

JAHN, H. A., and VAN WIERINGEN, H. (1951). *Proc. Roy. Soc., Ser. A* **209**, 502.

KLEIN, D. J., CARLISLE, C. H., and MATSEN, F. A. (1970). *Advan. Quantum Chem.* **5**, 219.

KOTANI, M., AMEMIYA, A., ISHIGURO, E., and KIMURA, T. (1955). "Tables of Molecular Integrals." Maruzen Co., Tokyo.

LITTLEWOOD, D. E. (1950). "The Theory of Group Characters and Matric Representations." Oxford Univ. Press, London and New York.

LÖWDIN, P. O. (1955). *Phys. Rev.* **97**, 1509.

LÖWDIN, P. O. (1960). *Rev. Mod. Phys.* **32**, 328.

LÖWDIN, P. O. (1964). *Rev. Mod. Phys.* **36**, 966.

LÖWDIN, P. O. (1967). *Rev. Mod. Phys.* **39**, 259.

MCINTOSH, H. V. (1960). *J. Math. Phys.* **1**, 453.

MCLACHLAN, A. D. (1960). *J. Chem. Phys.* **33**, 663.

MATSEN, F. A. (1964). *Advan. Quantum Chem.* **1**, 59.

MATSEN, F. A., CANTU, A. A., and POSHUSTA, R. D. (1966). *J. Phys. Chem.* **70**, 1958.

MATTHEISS, L. F. (1958). *Mass. Inst. Technol., Quart. Progr. Rep., Solid-State Mol. Theory Group* **30**, 43.

MATTHEISS, L. F. (1959). *Mass. Inst. Technol., Quart. Progr. Rep., Solid-State Mol. Theory Group* **34**, 58.

MILLER, K., and RUEDENBERG, K. (1968). *J. Chem. Phys.* **48**, 3414.

PAULING, L. (1933). *J. Chem. Phys.* **1**, 280.

PAUNCZ, R. (1967). "Alternant Molecular Orbital Method." Saunders, Philadelphia, Pennsylvania.

RUEDENBERG, K. (1971). *Phys. Rev. Lett.* **27**, 1105.

RUMER, G. (1932). *Nachr. Akad. Wiss. Goettingen, Math.-Phys. Kl.* p. 337.

RUTHERFORD, D. E. (1948). "Substitutional Analysis." Edinburgh Univ. Press, Edinburgh. (Facsimile Ed., Hafner, New York, 1968.)

SALMON, W. I., and RUEDENBERG, K. (1972a). *J. Chem. Phys.* **57**, 2776.

SALMON, W. I., and RUEDENBERG, K. (1972b). *J. Chem. Phys.* **57**, 2791.

SALMON, W. I., RUEDENBERG, K., and CHEUNG, L. M. (1972). *J. Chem. Phys.* **57**, 2787.

SERBER, R. (1934a). *Phys. Rev.* **45**, 461.

SERBER, R. (1934b). *J. Chem. Phys.* **2**, 697.

SLATER, J. C. (1931). *Phys. Rev.* **38**, 1109.

SULLIVAN, J. J. (1968). *J. Math. Phys.* **9**, 1369.

THRALL, R. M. (1941). *Duke Math. J.* **8**, 611.

WEYL, H. (1931). "Gruppentheorie und Quantenmechanik," 2. Auflage. Hirzel, Leipzig. [Engl. Ed., "The Theory of Groups and Quantum Mechanics" (H. P. Robertson, transl.). Dover, New York, 1950.]

WIGNER, E. P. (1931). "Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren." Vieweg, Braunschweig. [Expanded Ed., "Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra" (J. J. Griffen, transl.). Academic Press, New York, 1959.]

Yamanouchi, T. (1936). *Proc. Phys.-Math. Soc. Jap.* **18**, 623.
Yamanouchi, T. (1937). *Proc. Phys.-Math. Soc. Jap.* **19**, 436.
Yamanouchi, T. (1938). *Proc. Phys.-Math. Soc. Jap.* **20**, 547.
Young, A. (1901). *Proc. London Math. Soc.* **33**, 97.
Young, A. (1902). *Proc. London Math. Soc.* **34**, 361.
Young, A. (1928). *Proc. London Math. Soc.* **28**, 255.
Young, A. (1930). *Proc. London Math. Soc.* **31**, 253.
Young, A. (1932). *Proc. London Math. Soc.* **34**, 196.

Current Problems and Perspectives in the MO–LCAO Theory of Molecules

GIUSEPPE DEL RE
Cattedra di Chimica Teorica
Istituto di Chimica dell'Università
Napoli, Italy

I. Introduction

Recent quantum chemical research has placed much emphasis on the development of computational methods. It has been much less concerned with the physical significance of their results. This has created an acute need for critical analyses, as is shown by recent initiatives (Smith, 1972). The present study is intended to contribute to that critical work.

The main points that are under fire bring into play both the foundations of what has become a traditional approach of quantum chemistry—the MO–LCAO–CI scheme in its minimal-basis form, and the basic tenets of the general theory of molecules—atoms, bonds, and concepts derived therefrom. Some researchers claim that complete and reliable *ab initio* computations can be made for large molecules given enough time and money, that this rules out the need for simplified schemes, and that such things as bonds are only special features of simplified schemes.

The tendency to brush away most of what was built by chemistry is not easily acceptable both in view of the basic rule that working theories should not be discarded, but incorporated in more general ones, and in connection with the fact that the practical work of chemists is still based on bonds, atoms, and simple orbitals.

We wish to present here, as it were, the case for the MO–LCAO theory, meaning by the word theory a consistent if simplified interpretational scheme of a class of facts. Some aspects of that case are clear, some are not; the basis problem, as will be seen, is particularly serious and is still open; concepts like localization, hybridization, etc., are either defined in different ways or are rejected entirely as being too vague to be meaningful.

We shall examine in particular the significance of a one-electron scheme as a theoretical model for the interpretation of the behavior of molecules and the definition of the theoretical concepts of chemistry in connection with a quantum- mechanical model. This study is a review in the sense that it is intended to cover a rather extensive field, but it is not an attempt to cover the literature on the topics treated. It may even happen that some critical remarks are not referenced to the original papers; this is due to the fact that they are so to speak buried in essentially formal or applied papers, so that it is not easy to decide exactly where and by whom they were first made.

II. Interpretation vs. Computation

In the development of quantum chemistry there have been several largely overlapping but distinct phases. At the beginning the only problem was to show that the laws of quantum mechanics accounted for the chemical bond as well as for the spectrum of hydrogen. That work led to the two general schemes (VB and MO–LCAO) which later found both a number of successful applications and a formal organization. (c.f., Mulliken, 1932a,b, 1933, 1948, 1949, 1951, 1960, 1966a,b; Slater, 1931a,b, 1963; Van Vleck, 1933, 1934; Pauling, 1931).

The main goals of that work were (1) to apply quantum mechanics in order to account for the major features that make one molecule different from another, or at least to explain deviations from the simple atom–bond model, (2) to provide a rigorous framework within which theoretical quantities useful for interpreting chemical facts could be defined.

The greatest achievement was probably the explanation of conjugation, where a simple model was introduced to associate a quantum mechanical stationary-state picture with the purely chemical structural formula of a molecule, and then the basic concepts regarding coupling were used to show how conjugation arises. It is useful to summarize the argument without those details which are only essential in a rigorous formal discussion (for the latter, cf. McWeeny and Sutcliffe, 1969; Malrieu *et al.*, 1969).

A typical molecule can be treated as a system of independent electron pairs, one per bond. Therefore, a stationary state of a molecule (in particular its ground state) can be obtained as the product of states of individual pairs; the pair states depend only on the types of bonds, not on the particular molecule under study. In mathematical symbols, let ψ_0 denote the molecular ground state; then, if $\mathscr{A}$ is the antisymmetrization operator,

$$\psi_0 \approx \mathscr{A}(1_0)(2_0)\cdots(p_0)\cdots(n_0) \tag{1}$$

where n is the number of bonds, and (p_0) is the ground state of the pth bond (pair).

Equation (1) admittedly gives an approximate picture of ψ_0, but is supposed to be a rather good approximation whenever the given molecule obeys the additivity rules. However, it *must* break down when two pairs interact quite strongly and have very close groundstate energies. Then the four electrons of the two pairs must be treated as a whole, because the state where each pair is described as if it were alone is no longer even an approximate stationary state of the system. This is enough to explain conjugation—indeed is a redefinition of conjugation in quantum mechanical language—provided we define conditions for coupling that are fulfilled by pairs associated with conjugated double bonds. For this we decide that a necessary condition for independence is localization in different regions of space; then we divide pairs into σ and π pairs, the latter being somewhat less concentrated than the former; finally, we claim that parallel π pairs associated with next-neighbor bonds partly occupy the same region of space (i.e., their differential overlap is not zero). This gives us immediately the conclusion that conjugation is observed only with conjugated double bonds.

The above argument is qualitative, but it is easy to see in it the points that led to a shift from the conceptual and interpretational to the quantitative point of view. In fact, it defines implicitly two limiting cases, one where no conjugation exists and one where conjugation is full. One can associate to those limiting cases molecules like cyclohexene and benzene, but then the question arises of treating intermediate cases like naphthalene, pyridine, butadiene. There one can only reach useful conclusions by less qualitative arguments, and, as in most physical problems, one must define an *ad hoc* quantity which can vary continuously between the two limits, e.g., the famous (or infamous) resonance energy, and draw conclusions from more or less accurate numerical information (Del Re, 1971).

This explains how the second phase of recent quantum chemical work was entered; in particular it explains how the great interest in the Hückel method arose (Hückel, 1931; Coulson and Longuet-Higgins, 1947a,b, 1948; Pullman and Pullman, 1951). With this method the choice of the semiempirical parameters became a crucial point (Pullman and Pullman, 1960; Streitwieser, 1961).

Most of the work in this phase was done on conjugated π systems, because only conjugated molecules did not obey the additivity rules which held quite well for the other organic molecules; to many chemists, especially interested in predictions, the very validity of the additivity rules must have appeared a satisfactory theory. Actually, inductive effects are also introduced in organic chemistry to account for deviations from additivity rules of saturated (σ) bonds. Yet, with a few exceptions (Sandorfy, 1955; Del Re, 1958, 1964), the properties of saturated systems were not given as much attention. Indeed, the fact that the interest of experimentalists was also concentrated on unsaturated molecules led to great emphasis on the problem of predicting quantitatively specific properties of π systems rather than on the problem of explaining at least crudely the not less mysterious properties of saturated systems. This greatly contributed to make quantum chemistry, so to say, computationally oriented.

The point of view having changed, approximations had to be justified only in terms of convenience and ease of computation; it was not enough to claim that certain details did not change the essential lines of the picture. It is one thing to say that a certain quantity can be neglected or roughly estimated because this does not affect, say, conclusions regarding the behavior of the solutions of a certain differential equation, another thing is to say that it can be neglected because it does not affect seriously the quantitative reliability of the results. Moreover, in this new light it does seem quite unsatisfactory that certain parameters should be adjusted on experimental results to compensate for approximations.

The change in point of view coincided with a great development in the technology of computers, and this explains all that came later. After the improvements of semiempirical π-electron methods (Berthier, 1963) came the so-called *ab initio* calculations (Richards *et al.*, 1971).

Around 1960 the quantum chemists were divided into two camps: the pessimists, who thought that not over ten electrons could ever be treated in a rigorous way, and the optimists, who scorned semiempirical methods even for fairly large molecules and thought that all work should be concentrated in trying to draw quantitative predictions directly from the Schrödinger equation (Hall, 1959). The rapid development of computers seemed to suggest that the optimists were right, and extrapolation discouraged for a while semiempirical work.

When it became apparent that *ab initio* calculations would not automatically explain the properties of large molecules, the idea of coming back to approximations and models including σ electrons met with great favor, the more so because of the success among experimental chemists of rules and computations based on the "obsolete" semiempirical point of view (Gilchrist and Storr, 1972). The new rules actually supported the thesis that the qualitative physical point of view is very useful also in chemistry, but they led to a revival of semiempirical computations, emphasis remaining on quantitative predictions.

Thus came the third phase of quantum chemical research, the phase of all-valence-electrons methods for large molecules (Klopman and O'Leary, 1970; Kutzelnigg *et al.*, 1971, and references therein). These new methods do have a number of novel features in addition to including σ electrons, with respect to the old semiempirical methods; they do not try to adjust most of their parameters on experimental properties of selected molecules, but rather use atomic experimental properties and more or less "reasonable" numerical approximations for the various parameters [e.g., the Wolfsberg–Helmholz formula (Wolfsberg and Helmholz, 1952; Berthier *et al.*, 1966)]. Many logically unjustified and perhaps inconsistent assumptions are knowingly made in this connection on the grounds that good predictions are thus obtained. For instance, there appear overlap integrals which refer to Slater atomic orbitals, no reason for this choice being given. The most active members of the school of thought of the early 1970s even state explicitly that they do not think quantum chemistry should be but a tool for quantitative predictions of molecular properties (Dewar, 1963; Dewar and Haselbach, 1970).

In short, it would seem that the original problem of explaining molecular properties has been abandoned altogether.

Actually, we may well be at the beginning of a fourth phase, as is quite unobtrusively suggested by Wahl's (1966) density maps, one where clear-cut connection with rigorous formalisms and ability to explain facts are required of theoretical treatments even to the detriment of accuracy in quantitative predictions. This would mean going back to the old goal of quantum chemistry—interpreting and correlating facts in the light of quantum mechanical principles; it would mean to shift priority to the question *why?*, as opposed to the question *how much?*

This shift in priority does not concern *ab initio* computational methods, insofar as they were never conceived as anything but ways of providing the raw material for further analyses; it does affect both *ab initio* and semiempirical methods by requiring that the model with which they are associated should be suited to interpretational work.

The change in emphasis is just beginning; a critical revision of goals and concepts of quantum chemistry must be the first step. The foundations for this revision are ready. On one side, the lifelong work of researchers like Mulliken has kept alive the conceptual aspects of quantum chemical research; on the other side, the great computational effort devoted to MO–LCAO computations has led to a deep understanding of that method. It is mainly necessary to start putting known things together in a new critical context.

III. Problems of Theoretical Chemistry

Theoretical chemistry originated basically from the necessity of finding some criteria of classification for the immensely rich world of molecules. Now, to classify related objects with no immediately evident common characteristics one must go a long way toward building a "theory," as is known to everybody from the example of the theory of evolution. With molecules, classification required the concepts of atom, valency, and bond, and the notion of chemical formula. A distinction between different molecules was made by the hypothesis that the number and type of the atoms and the arrangement of the bonds was what made one molecule different from another.

Molecules have chemical and physical properties. The notion of bond was applied to systematize the former in statements as: " Molecules having a $C{=}O$ add HCN," the latter by rules like the additivity rules of energies and electric moments. With the development of structural studies, it was possible to associate with bonds also lengths and angles.

These three points constitute what one could call the theory of chemistry nonphysical, meaning that the notion of bond does not originate from any theory on the ultimate structure of matter. The stage that brought to the construction *ex novo* of concepts and laws useful for classification work had thus to be followed by a stage that demanded the derivation of those concepts and rules from the first principles of physics. The procedure that seemed to find general consensus after many years of exploratory work consisted in trying to solve the fundamental equations for a molecule and look for bonds in the solution. Unfortunately the analysis of molecular wavefunctions is so complicated that only a few serious efforts have been made so far to find the counterparts of the chemical atom and of the chemical bond in molecular quantum states, and the outcome of some of them is paradoxically the claim that the bond concept is not a uniquely defined one (Clementi and Popkie, 1972a; Clementi, 1967). Yet, a bond is not only the foundation of additive rules and stereochemistry, but a fundamental concept related to the notion of valency and of chemical formula. In addition to approximate geometries, heats of formation, electric dipole moments, etc., even the number of isomers of a given saturated hydrocarbon can be determined using the bond concept.

The major attempts to fill the gap between quantum chemistry and the classical theory of chemistry are associated with the density matrix concept (Löwdin, 1955; McWeeny, 1954, 1960). The idea of looking at the chemical formula as a scheme representing a partitioning of the electron distribution was extensively discussed by Ruedenberg (1962). However, the realization that theoretical partitionings are not unique casts a shadow of mistrust over certain definitions. Ways of avoiding ambiguities were suggested, although the times were not ripe for attracting too much attention to them. One was the pair-function scheme, which went back to the Lewis idea of bonds representing electron pairs, and there exists a full computational scheme based on it (McWeeny and Sutcliffe, 1969; Malrieu *et al.*, 1969). Another was the idea of block factorizing as far as possible, by means of hybridization, a given one-electron Hamiltonian (Del Re, 1963; Veillard and Del Re, 1964; Berthier *et al.*, 1964), and the later version of that idea, according to which the factorization was imposed on the density matrix (McWeeny and Del Re, 1968). Finally, a definition of a bond corresponding to a procedure strictly parallel to that which led historically to the bond concept was also suggested. The idea consisted in associating to bonds transferable localized two-center two-electron wavefunctions, whose products should match as closely as possible the "exact" wavefunctions of molecules containing the given bonds (Nelander and Del Re, 1970).

Until very recent work (Clementi and Popkie, 1972a; England and Gordon, 1971), the attempts mentioned above remained almost a dead letter mainly because of the great popularity of straightforward computations on larger and larger molecules. Therefore, by and large the situation is not clearer today than it was twenty years ago; there remain many unsettled conceptual and interpretational problems in quantum chemistry, which all stem from the fact that neither the old concept of bond, nor new concepts like hybridization, delocalization, etc. have been fitted into a consistent picture. Indeed, both old and new concepts have even been dismissed by the most recent semiempirical methods—which nevertheless take the chemical formula as an input datum—on the grounds that they are not useful for computational purposes.

The real reason for that dismissal is probably that those concepts are difficult to assess and use in the absence of general consent on their definitions.

An attempt to improve the situation in this connection can be based on the MO–LCAO–CI scheme (Mulliken, 1949). This means that one has to treat first of all the so-called basis problem. In fact the idea that evolved out of the work of the pioneers (Mulliken, 1932a,b, 1933, 1948, 1949, 1951, 1955, 1960, 1966a.b; Slater, 1931a,b, 1963; Van Vleck, 1933, 1934; Pauling, 1931; Hückel, 1931; Coulson and Longuet-Higgins, 1947a,b, 1948; Pullman and Pullman, 1951; Pauling, 1945; cf. Del Re, 1967) was that one could identify the chemical atom with a limited set of atomic orbitals entering one-electron molecular wavefunctions, and that hybridization and localization would fit in that scheme. Use of extended bases and "proofs" that localization was an arbitrary process led to much mistrust of that idea on the part of many specialists, even though that scheme is now part of the elementary theory of chemistry and has proved to be of great heuristic value.

IV. The Basis Problem

The present form of quantum chemistry is largely the outcome of the introduction of the concepts of atomic orbital (AO), hybrid atomic orbital (HAO), and molecular orbital obtained by linear combination of atomic orbitals (MO–LCAO). All were proposed as a result of physical considerations (Slater, 1963; Pauling, 1945). For instance, the aufbau principle of atomic quantum mechanics was used to translate the octet rule into an acceptable quantum mechanical language. A physical model underlies that translation: an atom in a molecule was treated as if it were an isolated atom capable of filling its orbitals with the electrons of neighboring atoms. To rationalize this picture the present MO–LCAO–CI scheme was gradually developed, based on the Ritz variational method (Mulliken, 1949; Kemble, 1958). Unfortunately, as a by-product of the same process, the atomic orbitals came to be interpreted as more or less arbitrary basis functions. Eventually, with the introduction of Gaussian orbitals, the name MO–LCAO came to mean only a computational scheme involving a one-electron model.

Nevertheless, many interpretational considerations are still made on the assumption that the details of the (minimal basis) MO–LCAO–CI scheme describe physical aspects of molecules. For instance, lone pairs, overlap, etc., belong to the current language of chemistry. The fact that the corresponding picture is not particularly good for accurate energy computations is irrelevant for many applications. Therefore, a discussion of the problem of the choice of AO bases in connection with their physical meaning is in order. We shall use systematically the idea that a physical meaning is given to something whenever a (static or dynamic) physical model can be associated with it.

The general basis problem of quantum mechanics can be stated at a given instant of time in a vector space usually with an infinite number of dimensions (the Hilbert space).

This postulate means that it is possible to give (in infinitely many ways) a basis, i.e., a set of linearly independent "states" which can be taken discrete for simplicity, forming a row matrix

$$|\chi\rangle = (|\chi_1\rangle|\chi_2\rangle \cdots |\chi_\mu\rangle \cdots) \tag{2}$$

such that any given state of the given system is a linear combination of them

$$|\psi\rangle = \sum_\mu C_{\mu j}|\chi_\mu\rangle \tag{3}$$

or

$$|\psi\rangle = |\chi\rangle \mathbf{C}_j \tag{4}$$

where the symbol $\mathbf{C}_j$ denotes a column vector which can be considered as the jth column of a matrix $\mathbf{C}$.

The mathematical counterpart of the physical superposition principle is the statement that one can always find sets of functions (say, in configuration space) belonging to a suitably defined class (e.g., see Kemble, 1958) such that any state of a physical system is represented (again in configuration space) by a linear combination of the functions in question.

In virtue of the variational principle (the Ritz method) and of intuitive considerations, one can try to obtain an approximate expression of certain states of a given system by taking only a finite number n of elements of a suitably chosen basis. In particular, the variational principle insures that, if the state under consideration is the ground state of the given system, there exists an optimal choice of the elements of $\mathbf{C}_j$, and, with certain reservations, the method giving $\mathbf{C}_j$ can give a reasonable approximation of the other states.

For our scope, it is sufficient to state that given a finite ("truncated") basis $|\chi\rangle$ consisting of n elements, one can obtain n "best" approximate stationary states by inserting in Eq. (3) the columns of the matrix **C** which fulfills the equation

$$\mathbf{HC} = \mathbf{SCE}, \qquad \mathbf{C}^{\dagger}\mathbf{SC} = \mathbf{I}, \tag{5}$$

where the dagger means the Hermitian conjugate, **H** is the matrix representation of the Hamiltonian operator $\mathscr{H}$ of the system in the truncated basis $|\chi\rangle$, and **S** is the corresponding metric matrix

$$\mathbf{S} = \langle\chi|\chi\rangle; \tag{6}$$

E is a diagonal matrix which gives the approximate energies of the approximate states defined by Eqs. (5) and (3), and **I** is the unit matrix.

Let us grant that Eq. (5) can give a reasonable approximation if $|\chi\rangle$ is properly chosen. How do we extract $|\chi\rangle$ from (an arbitrary) $|\chi\rangle_\infty$? A current technique, which many (against authoritative warning, Löwdin, 1956) consider foolproof consists in selecting a computationally reasonable basis and increasing the number of elements so as to obtain convergency at least for the ground state. This technique is acceptable, but gives no assurance of actual correspondence between the limit obtained and the true limit (for certain important elements could have been systematically neglected), nor is it physically significant, because the basis is chosen only on grounds of computational convenience.

The explanation of differences between molecules in terms of LCAO molecular orbitals requires that the basis atomic orbitals should be limited in number to a minimum corresponding to the main characteristics of the molecule. For instance the customary habit of assigning four orbitals to each atom of the first row of the periodic table corresponds to a criterion of simplicity which is essential to chemists. On the other hand, it is well known that so far calculations carried out over a limited basis (Richards *et al.*, 1971; Ransil, 1963) do not give satisfactory results from the point of view of energies. A question related to physical interpretation is thus whether it is possible to reconcile the use of an extended basis with an analysis in terms of a small number of atomic orbitals (Mulliken, 1966a). This question could be discussed *a posteriori*; in our view it is better to discuss it without strict reference to calculations satisfying special criteria of accuracy. In this connection it may be important to warn against the tendency to identify atomic orbitals with Slater orbitals (cf. Clementi and Raimondi, 1963). First of all, one can replace the latter by analytical approximations of Hartree–Fock free-atom orbitals satisfying special optimization conditions (Cusachs and Aldrich, 1971) or by other sorts of orbitals still satisfying a variational principle for energy (Rastelli and Del Re, 1969). In general, one must consider as orbitals functions centered on the various atoms of a molecule [Mulliken's MAO's (Mulliken, 1966a)] for which one can still speak of s, p. d, ... symmetry without implying any special form of the radial parts (Del Re, 1967); the latter should be chosen, for instance, with the aim of obtaining from Eq. (5) an approximate wavefunction whose partitioning in terms of those atomic orbitals would satisfy particular conditions required by the interpretation scheme placed at the basis of the calculation (e.g., localization of bonds, orthogonality, localization of charges, etc.).

In accordance with the above remarks, consider a truncated basis of functions which are atomic orbitals in the sense specified above. In order to fulfill the condition that this basis could in principle be extended to give any desired accuracy to energy calculations, one must assume that the finite set of orbitals

$$|\chi\rangle = (|\chi_{1A_1}\rangle, |\chi_{2A_1}\rangle, \ldots; |\chi_{1A_2}\rangle, |\chi_{2A_2}\rangle, \ldots) \tag{7}$$

belonging to the N atoms A_1, A_2, ..., A_N of the given molecule does belong to a complete set $|\chi\rangle_\infty$. The set $|\chi\rangle$ is thus assumed to be formed by N subsets $|\chi_A\rangle$ each relating to one atom A_j and formed by n_j elements.

Let us introduce also the sets $|\bar{\chi}\rangle$ and $|\bar{\chi}\rangle_\infty$ whose elements are Slater orbitals (or other atomic orbitals whose form is fully specified). We can obtain the basis $|\chi\rangle$ by transformation in the Hilbert space spanned by $|\bar{\chi}\rangle_\infty$ followed by (a) truncation of the basis, (b) a linear transformation of $|\bar{\chi}\rangle$.

It is easily understood that the two types of transformation are not equivalent in general. In fact if $\mathbf{\Pi}$ denotes the projection operator which singles out $|\chi\rangle$ from $|\chi\rangle_\infty$, and $\mathbf{T}$ and $\mathbf{T}_1$ are two linear transformations, the two cases correspond to

$$|\chi\rangle = |\bar{\chi}\rangle_\infty \mathbf{T}\mathbf{\Pi}, \tag{7a}$$

$$|\chi\rangle = |\bar{\chi}\rangle_\infty \mathbf{\Pi}\mathbf{T}_1. \tag{7b}$$

Therefore, unless $\mathbf{T}$ and $\mathbf{T}_1$ are related in a very special way, cases (a) and (b) are not equivalent. From the practical point of view, case (a) relates to the problem "choice of the form of the atomic orbitals," whereas case (b) relates to the problem "choice of preliminary combinations of atomic orbitals already forming a truncated set." Of course one can define a single transformation leading from $|\bar{\chi}\rangle_\infty$ to $|\chi\rangle$,

$$|\chi\rangle = |\bar{\chi}\rangle_\infty \mathbf{T}\mathbf{\Pi}\mathbf{T}_1, \tag{8}$$

which can be read: We choose a complete set in the Hilbert space associated with the operator we are interested in; we combine its elements so as to bring in the first places certain functions we are interested in; we single those functions out of the whole complete set; we combine them linearly.

More precisely, we can think of our truncated atomic orbital set as being obtained by building atomic orbitals of the types *s*, *p*, *d*, etc., which are linear combinations of Slater orbitals of the same species but with different principal quantum numbers, and by hybridizing them, i.e., combining them within subsets corresponding to individual atoms.

Before discussing more systematically the question of physical significance, it is important to mention something about completeness and overcompleteness of the set $|\chi\rangle_\infty$ to which our set $|\chi\rangle$ ideally belongs. This question of overcompleteness was emphasized by Löwdin (1956), and is physically very important because, for high quantum numbers, the atomic orbitals evidently become so large that the very idea of atomic orbitals localized on the individual atoms loses its meaning; to an electron located very far from the molecule, the latter appears very much like the corresponding united atom. This consideration suggests that our set $|\chi\rangle$ should be completed not necessarily by more atomic orbitals centered on the different atoms, but by a set of appropriately chosen one-center functions. Thus, in order to extend our basis, we should not think of extending the individual atomic orbital sets $|\chi_A\rangle$ to include all the possible orbitals of every atom, but rather proceed as follows. We take first $\sum_{j=1}^{N} n_j = N_{\text{tot}}$ orbitals; then we consider the elements of, say, a hydrogen-like set $|\phi\rangle$ centered at a particular point of the given molecule. Finally we build a matrix $\mathbf{D}$ whose first N_{tot} columns are zero, and whose remaining columns are normalized, orthogonal to one another, and orthogonal to the rows of the (rectangular) matrix whose elements are the overlap integrals between the N_{tot} elements of $|\chi\rangle$ and the elements (infinite in number) of $|\phi\rangle$. Then the set

$$|\phi\rangle_{\text{tot}} = (|\chi\rangle + |\phi\rangle\mathbf{D}) \tag{9}$$

is complete, entirely equivalent to $|\phi\rangle$, and has as its first N_{tot} elements those of $|\chi\rangle$.

On the basis of the above considerations we can carry out our discussion keeping in mind that it will always be possible, at least in principle, to improve a calculation by introducing overall effects through the inclusion of elements of the set $|\phi\rangle\mathbf{D}$. In principle one could even pass gradually to the united atom orbitals by adding to $|\chi\rangle$ first some orbitals centered at points different from and fewer in number than the nuclei forming the framework of the molecule, then more orbitals centered at an even smaller number of points, and so on.

Consider now the set $|\chi\rangle$. What are the conditions one may wish to impose on these orbitals in order to "prepare" the molecular orbitals and/or the CI wavefunctions for a physical analysis?

Of course, accuracy in predicted energies is not a good criterion, for, as is well known, it is by no means certain that a limited molecular calculation carried out with the sole aim of obtaining a "good" energy would yield a reliable description of a molecule in the chemical sense. In fact, minimal basis molecular calculations using Slater orbitals with a variable orbital exponent (Ransil, 1963; McLean and Yoshimine, 1967) lead to orbital exponents close to those of the Slater orbitals of the free atom, but this is not necessarily an indication that Slater orbitals are indeed "good" atomic orbitals. For one thing, use of non-Slater AO's might lead to worse energies, but, say, to much better relative charge distributions, and one could always play with the orbital exponents of the Slater orbitals entering the expression of a general atomic orbital so as to obtain an energy close to or equal to that of the corresponding Slater orbital (Rastelli and Del Re, 1969).

The above considerations suggest that the only reasonable way of tackling the choice of a truncated basis consists in defining *models and limiting cases* to be associated with it (as well as with other approximations possibly made in writing Eq. (5) explicitly). This is a procedure extensively accepted by molecular spectroscopists (in particular Salahub and Sandorfy, 1971) and adopted systematically by Mulliken. For instance, the latter introduced the MAO concept by the following physical ideal process of formation and dissociation of H_2^+ in the MO scheme (Mulliken, 1966a) when the internuclear distance goes from zero to very large values: United Atom Orbital (UAO) → slightly modified United Atom Orbital (mUAO) → much Modified United Atom Orbital (MUAO) or Linear Combination of Modified Atomic Orbitals (LCMAO) → Linear Combination of slightly modified Atomic Orbitals (LCmAO) → Linear Combination of Atom Orbitals (LCAO).

This scheme requires careful analysis for large internuclear distances in many-electron molecules because of the well-known failure of the unsophisticated LCAO scheme to describe dissociation. But it is important for us as a very pertinent example of a model (H_2^+) based on a physical limiting case (the united atom). We examine now the notion of a model in more detail. Another explicit definition of a model by a quantum chemist is by Pople (in Smith, 1972) who called a model a semi-empirical scheme with a number of unjustified simplifications. This is a very special interpretation of the model notion, as Löwdin had pointed out to the present author several years before Pople's remarks, and is not the same as the physicists' interpretation (Andronov *et al.*, 1966).

An interesting version of a quantum mechanical model explicitly proposed and discussed as such has been recently developed by Hall (1973), who replaces the charge density by a point charge distribution preserving certain observables.

V. Model Treatments of Molecules

A simplified treatment based on a model involves replacement of an actual physical system (say, a molecule) by a simpler one (the model) which is treated in quite a rigorous way. The fact that approximations are introduced in constructing the model rather than in the mathematical development implies that the various aspects of the corresponding treatment can be given a physical interpretation.

The treatment of a molecule in terms of a model reduces to replacing the true Hamiltonian of the molecule by one where certain details are neglected, *viz.* by some sort of idealization. For instance, to imagine that the 1*s* electrons of the atoms of a molecule are concentrated in the nuclei corresponds to is a model treatment. The independent-particle approximation of a molecule is a model if it involves a well-defined Hamiltonian.

At first sight, the "basis problem" of molecular calculations is conceptually different from the problem of finding suitable models, because it involves the choice of a subspace of the Hilbert space in which a physically significant treatment of a molecule can be carried out. But we can easily see that, under certain conditions, the choice of the basis is equivalent to the choice of a model.

In quantum mechanics, a physical system is completely defined if one knows either (a) the particles of which it consists and their interactions, or (b) the energy levels, classified according to other observables in case of degeneracy, together with complete information regarding their coupling if they correspond to quasi-stationary states. Therefore, one type of model (which we shall call type A) consists in neglecting or averaging certain interactions between particles (say, the independent-particle model of a several-body system). The other type B consists in replacing the actual energy-level spectrum by a simpler one. This is done, for instance, in the two-state model of ammonia used in the theory of the maser (Feynman 1966). A composite model is possible, and we shall call that type of model AB.

The choice of a basis may correspond to the choice of a type A reference model, for the basis elements may be defined as the time-independent parts of the stationary states of a physical system having the same Hilbert space as the given system, and then evidently a type A model has been chosen. Truncation corresponds to a type B model. In fact, truncation of the basis just means that states expressed in terms of the truncated basis are not coupled by the given Hamiltonian with states expressed in terms of the part of the basis which has been left out by the truncation. In other words, the basis is separated into two subsets, one a finite set $|\chi\rangle$ consisting of m elements, the other the complementary set $|\chi'\rangle$

$$|\chi\rangle_\infty = |\chi\rangle + |\chi'\rangle \qquad \text{[cf. Eq. 9]}, \tag{10}$$

and we assume that the basis and the Hamiltonian operator $\mathscr{H}_{\text{model}}$ are such that

$$\langle\chi|\mathscr{H}_{\text{model}}|\chi'\rangle = \langle\chi'|\mathscr{H}_{\text{model}}|\chi\rangle = 0. \tag{11}$$

This shows that the choice of a physically significant truncated basis is equivalent to the choice of an AB type model, for it is the result of (a) the definition of a special complete basis and (b) the assumption that the set $|\chi\rangle$ of m basis elements are separated from the rest because of negligible coupling. It also shows how the choice of the basis is related to the choice of a model and of the corresponding Hamiltonian (O'Malley and Geltman, 1965).

The above remarks apply to truncated CI bases as well as to AO bases. As we concentrate on the latter, $|\chi\rangle$ will represent an AO basis. Finite basis MO–LCAO treatments are based on a model derived from an independent particle model (IPM type A). In particular a typical semiempirical method of that kind assumes that: (1) The molecule can be treated as an "IPM molecule," i.e., as a system consisting of n noninteracting electrons moving in the field of an effective core, electron–electron interactions being taken into account by appropriate contributions to the core potential; (2) there exists a finite set $|\chi\rangle$ of atomic orbitals such that a certain set $|\psi\rangle$ of n states of an electron in the IPM molecule may be represented as a linear combination of them.

Assumption (2) is expressed by Eq. (11) and demands the remarks made above on type B models because a finite set of atomic orbitals can only give a finite set of (linearly independent, say orthogonal) molecular orbitals. This means that an infinite number of states of the molecule [idealized according to assumption (1)] are not coupled with those belonging to $|\chi\rangle$.

The algebraic procedure for determining the MAO's defined in the preceding section and in Rastelli and Del Re (1969; Del Re, 1967, 1969) from the general conditions (11) can be specified as follows. Let $|\mathbf{s}_{A_1}\rangle$, $|\mathbf{s}_{A_2}\rangle, \ldots$ denote the subsets formed by STO's (Slater orbitals) 1s, 2s, ... on atoms A_1, A_2, etc., and $|\mathbf{p}_{A_1}\rangle$, $|\mathbf{p}_{A_2}\rangle, \ldots$ denote the corresponding p subsets. Then the MAO minimal basis can be written

$$|\chi\rangle = (|\mathbf{s}_{A_1}\rangle|\mathbf{p}_{A_1}\rangle \cdots |\mathbf{s}_{A_2}\rangle|\mathbf{p}_{A_2}\rangle\cdots)\begin{pmatrix} \mathbf{Q}_{A_1}^{(s)} & & & & \\ & \mathbf{Q}_{A_1}^{(p)} & & & \\ & & \ddots & & \\ & & & \mathbf{Q}_{A_2}^{(s)} & \\ & & & & \ddots \end{pmatrix} \tag{12}$$

where a $\mathbf{Q}_{A_j}^{(l)}$ matrix ($l = s, p, d, \ldots, j = 1, 2, \ldots$) has as many rows as the elements of $|\mathbf{l}_{A_j}\rangle$ and as many columns as the number of orbitals of the given symmetry type (l) in the valence basis of A_j.

Introducing the one-center complementary set $|\boldsymbol{\phi}\rangle\mathbf{D} = |\chi'\rangle$ of Eq. (9), we have the orthogonality condition

$$\mathbf{D}^{+}\langle\boldsymbol{\phi}|\mathbf{l}_{A_j}\rangle\mathbf{Q}_{A_j}^{(l)} = \mathbf{0} \tag{13}$$

for all l and j, and the separation condition

$$\mathbf{D}^{+}\langle\boldsymbol{\phi}|\mathscr{H}_{\text{model}}|\mathbf{l}_{A_j}\rangle\mathbf{Q}_{A_j}^{(l)} = \mathbf{0}. \tag{14}$$

[Note that Eq. (13) is an alternative to making $|\boldsymbol{\phi}\rangle$ orthogonal to $|\chi\rangle$, as assumed in Eq. (9).]

The existence problem for the solutions of Eqs. (13) and (14) is rather complicated and largely hinges on specific circumstances. In general most elements of $|\boldsymbol{\phi}\rangle$ will give negligible contributions both in Eqs. (13) and (14); the remaining nonvanishing terms will give a finite number of linear equations in the unknowns $\mathbf{Q}_{A_j}^{(l)}$, which can certainly be solved by taking enough STO's of different quantum numbers for every set $|\mathbf{l}_{A_j}\rangle$. Even if matters are thus simplified, handling the orbital exponents may represent a serious difficulty when optimizing with respect to the energy, because the $\mathbf{Q}$'s will depend on them.

The fact that in many cases we are only interested in states belonging to the Hilbert subspace spanned by $|\chi\rangle$ does not mean, of course, that the orbitals described by the set $|\chi'\rangle$ complementary to $|\chi\rangle$ are not interesting. On the contrary, they include the remarkable Rydberg states of molecules (Salahub and Sandorfy, 1971).

The classification of MO's into Rydberg and non-Rydberg involves, in addition to the independent-particle model, a one-center reference model of the sort suggested by Mulliken (see end of preceding section). The latter serves to provide the one-center part of the basis, $|\chi'\rangle$, while molecular states are approximated in the usual valence-shell orbital basis $|\chi\rangle$. The important point is that, even after separation according to Eqs. (13) and (14), certain MO's built with $|\chi'\rangle$ may give excited states lower in energy than the highest excited states built with $|\chi\rangle$. Thus some of the lower excited states could be either entirely described by electron distributions closely concentrated along the molecular frame or by atom-like distributions. The latter, with the rest of the excited states involving $|\chi'\rangle$, belong to sequences of the classical Rydberg type, and may be called Rydberg states.

The explicit formulas (13) and (14) show that the separation (11), if it is possible at all, certainly takes place only for a particular basis set. Therefore, assumption (2) is essentially a criterion for choosing the AO basis $|\chi\rangle$; it provides an explicit statement of the fact that the choice of $|\chi\rangle$ is not a mere matter of computational convenience even in the so-called *ab initio* treatments. In treatments based on " semiempirical models " that condition (and the choice of $|\chi\rangle$ in general) is vital, because the basis is supposed to contain only a very limited number of elements.

Equations (13) and (14) provide an explicit procedure for determining MAO's if a suitable model Hamiltonian is given and if $|\phi\rangle$ is chosen in a specific way. Pending that, the only clear conclusions we can state are: (a) One should expect that the atomic orbitals to which the elementary theory of chemistry refers differ from free-atom orbitals (FAO's); (b) simple modification of the orbital exponents of Slater type FAO's is not enough to generate suitable MAO's; (c) a definition of the atom *in situ* consistent with an MO–LCAO scheme must be the starting point of an analysis of the whole question (Mulliken, 1949).

Conclusion (a) is just a negative one. As is well known, the set of the FAO's of each atom is complete. Therefore, one can always express a molecular orbital as a superposition of the free-atom orbitals of a single atom; in a very broad sense, a molecular orbital is thus a special case of atomic orbitals; indeed accurate *ab initio* molecular calculations have been made in one-center expansions (Moccia, 1964). This shows that to determine explicitly the features that make an orbital atomic is not a trivial task.

The best way to solve this problem is to find a type A reference model whose stationary states are precisely the required MAO's, freedom being left to the latter for the conditions (11).

VI. A Hamiltonian for a Reference Model

Objections may arise concerning the necessity of introducing models to which a Hamiltonian (and hence a wavefunction) is associated. For instance Ruedenberg (1962) tried to define his valence and promotion atomic states in terms of densities and pair densities only. Clementi proceeded in a very similar way on the basis of a partitioning of energy (Clementi and Routh, 1972). The main reason why this is not entirely satisfactory is that inconsistencies may thus be allowed because the given model is not automatically extended to excited states. Moreover, to have a reference model in terms of potential energy operators and wavefunctions certainly enhances the usefulness of the model for interpretation. We emphasize also that the validity of a given model Hamiltonian is not dependent on how close its eigenvectors are to the "true" eigenstates, but on the ease of interpretation of the differences between the former and the latter.

It may be objected that it is not possible to give a model Hamiltonian which is clearly connected with the complete Hamiltonian. A mathematical analysis of this point is given in the Appendix. It is shown there that the general Hamiltonian of a molecule can be written in such a way that any one-electron Hamiltonian, in particular one given in matrix form, can be introduced in it and treated as the one-electron part of the whole molecular Hamiltonian.

With these premises, we focus our attention on the model that could provide as eigenstates the *in situ* atomic orbitals or MAO's.

The very attribute "atomic" indicates that the reference model should somehow correspond to a fictitious isolated atom. This is in contrast with the idea that MAO's are atomic orbitals *in situ*, namely in a situation where the individual nuclei are fairly close to one another. The only way of reconciling the two points consists in seeing whether, by neglecting certain coupling terms, one can treat a molecular system as a collection of isolated atoms. In principle there should be no difficulty with this problem, which requires just completion of Mulliken's scheme (end of Section IV). For infinite separation of the nuclei the states of one electron in the field of many nuclei must become those of an electron in the presence of a single nucleus.

Then, the states of the system formed by N (effective) nuclei and one electron when the nuclei are very far apart are: N states corresponding to the electron on each of the lowest energy atomic orbitals, N states corresponding to the electron in each of the first excited orbitals associated with the individual nuclei, and so on.

In operator form, the Hamiltonian that gives these states is the same effective one-electron Hamiltonian as that which is associated with the actual molecule. The fact that the special "atomic" states in question are eigenfunctions of that Hamiltonian stems from the fact that certain matrix elements between atomic orbitals vanish for large internuclear separation.

The approach suggested by these considerations consists in selecting an effective one-electron Hamiltonian for the given molecule, building the corresponding matrix over atomic orbitals containing some flexibility, neglecting all the contributions to the various matrix elements which come from atoms other than that to which every given atomic orbital is related, and minimizing the energy associated with each "atomic" Hamiltonian thus constructed by varying the flexible parameters of the atomic orbitals.

This procedure is not equivalent to ignoring that the atoms under study participate in molecules, because we start from an effective Hamiltonian where the distribution of all the electrons but one is assumed to be that of the actual molecule. The "isolated" atom to which every atomic orbital refers is not the free atom. We are imagining a process in which the atoms are slowly separated, starting from the situation in the actual molecule, without any modification of the electron distribution, thus defining something very close to Ruedenberg's (1962) promotion.

If the initial Hamiltonian is an effective one, the part of it which corresponds to a single standard *in situ* atom describes a hydrogen-like atom or a more complicated situation according to whether the average electron cloud is supposed to be concentrated in the nucleus or not. In the former case the atomic orbitals will show the typical degeneracies of a hydrogen atom. Therefore, they will be worse representations of the real situation; nevertheless, they will lead to a consistent scheme. In fact, hybridization will then appear as an obvious zero-order step in the preparation of the effective atoms for the formation of the bonds, and thus the process of construction of the molecular wavefunction from the atomic orbitals taken as building blocks will have the elegant features of the perturbation treatment.

The idea of a minimal valence AO or MAO basis, the question of physical significance, and the notion of a model or reference system have thus led to the conclusion that a separated-atom model Hamiltonian is what one must define in order to give a clear-cut meaning to a minimal valence basis of MAO's corresponding to the current sense of elementary chemistry. The separated atoms thus introduced may or may not be hydrogen-like atoms, but certainly they will not be the same as the corresponding free atoms.

Reasons for preferring hydrogen-like reference atoms can be given;

among others the general tendency to increase of orbital exponents, *viz.* contraction of orbitals. The main point, however, is that such "promoted" atoms must be in a one-to-one correspondence with atoms of elements in the typical valence states of classical chemistry. For only in that way can they fulfill the condition of rigorously embodying the chemical idea of an atom (which, it should not be forgotten, is the original form in which the modern atom first appeared in contemporary science). In other words, these *promoted* atoms must either be independent of the type of molecule in which they participate or—which is conceptually the same—change according to well-specified *a priori* rules.

To see the problem in practice, consider a reference Hamiltonian formally corresponding to one electron in the field of N effective point charges $Z_X{}^*$ (atomic units):

$$\mathscr{H}_{\text{model}} = -\tfrac{1}{2}\nabla^2 - \sum_X Z_X{}^*/r_X \tag{15}$$

r_X being the distance of the electron from the charge located at X. The possible choices of the effective charges Z^* can be discussed (Simon and Pujol, 1969) starting from the classical expression for the average effective nuclear potential created by N nuclei X and by n electrons

$$V_N = -\sum_X \frac{Z_X}{r_X} + \sum_k \bar{n}_k \int \frac{|\phi_k(2)|^2}{r_{12}}\, dv_2 \tag{16}$$

where k refers to MO's ϕ_k and the $\bar{n}_k$'s are effective occupation numbers such that their sum is $n - 1$. It is possible to write V_N in a form where a term of the point-charge type appears explicitly by introducing a partition based on the Mulliken approximation for orbital products:

$$\int \frac{|\phi_k(2)|^2}{r_{12}}\, dv_2 = \sum_{X,\xi} \frac{Q_{X\xi}^{(k)} v_{X\xi}}{r_X} + \text{correction}, \tag{17}$$

where

$$v_{X\xi} = r_X \int \frac{|\chi_{X\xi}(2)|^2}{r_{12}}\, dv_2 \tag{18}$$

is a "screening factor",

$$Q_{X\xi}^{(k)} = \operatorname{Re} \sum_k \sum_{Y,\eta} c_{X\xi,k}^* c_{Y\eta,k} S_{X\xi Y\eta} \tag{19}$$

is the Mulliken gross population [Mulliken (1955)] of the ξth AO of X for the kth MO; $c_{X\xi,k}$ is the coefficient of the ξth AO of atom X in the

kth MO; $S_{X\xi Y_\eta}$ is the overlap integral between the ξth AO of atom X and the ηth AO of atom Y.

The definition of $Z_X{}^*$ can be based on substitution of Eq. (17) in Eq. (16)

$$Z_X{}^* = Z_X - \sum_\xi \sum_k Q_{X\xi}^{(k)} \, v_{X\xi} \, \bar{n}_k \tag{20}$$

where it is necessary (a) to eliminate the dependence on the particular molecule under study as represented by the occupation numbers and by the orbital coefficients appearing in Eq. (19), and (b) to decide how the screening factors can be evaluated and/or the rules they must obey. The best recipe is probably one where a reference molecule is chosen as one where the atoms exhibit their standard MAO's; such a choice corresponds to the typical approach of chemistry to the classification of molecules. It does not settle the problem, however, because one must also define the molecular state from which the occupation numbers are taken and the forms of the screening factors (cf. Hall, 1954).

A complete assessment of the model of Eq. (15) has not been attempted so far, and the numerical data for such an assessment are not available. A convenient starting point may be found in the recent work by Clementi and Popkie, where an energy and population partitioning of methane, ethane, ethylene, acetylene is proposed and carried out. That is by no means a partitioning that recognizes the fundamental importance of having a clear-cut model in mind beforehand (indeed the authors equate physical significance with numerical accuracy), but it can serve as a basis for a preliminary evaluation of the Z^*'s, either directly or via the orbital energies (Clementi and Popkie, 1972a,b). For instance, one may use the valence-shell orbital energy of methane to obtain a crude estimate of the effective nuclear charge of a hydrogen-like carbon atom (nucleus plus 1*s* electrons) surrounded tetrahedrically by four partners—and, of course, the effective charges of the hydrogen atoms in the molecular environment. This gives 2.74 for carbon, 1.04 for hydrogen. Of course, the variational theorem applied to pure hydrogen-like orbitals must give the orbital exponents in a direct way.

Either by sticking to the model (15) or using a more flexible form of effective potential much can be done to improve the determination of the MAO's. The inclusion of some flexibility in the radial parts may give a much more sophisticated choice. For instance, one can replace the powers of the radius appearing in the radial parts of Slater orbitals by higher degree polynomials, so as to introduce "radial distortion" in addition to "scaling" (change in orbital exponent). This is strictly an atom *in situ* effect, for it is strictly zero in free atoms (Rastelli and Del Re, 1969).

The importance of suitable scaling in connection with *in situ* MAO's has been briefly reviewed and stressed in very recent work (Matheu and Ellison, 1972). The authors write: "The basis AO's for each atom are defined as eigenfunctions of a prescribed atomic Hartree–Fock Hamiltonian operator; however, the corresponding eigenvalues are made consistent with atomic integrals determined by calibration of a *semiempirical scaled atomic orbital* (SESAO) model to experimental atomic spectra." The Hartree–Fock atomic operator in question is obtained from the well-known Hartree–Fock equation by introducing a scaling factor s_X:

$$\mathscr{H}_{\text{model}} = -\tfrac{1}{2}\nabla^2 + s_X{}^2 V'_{X\xi} \tag{21}$$

where $V'_{X\xi}$ is the *scaled* form of

$$V_{X\xi} = -\frac{Z_X}{r_X} + \frac{1}{2} n_{0X\xi}(n_{0X\xi} - 1)J_{X\xi} + \sum_{\xi' \neq \xi} n_{0X\xi'}\left(J_{X\xi'} - \frac{1}{2} K_{X\xi'}\right) \tag{22}$$

(or of a modified expression designed to preserve symmetrical treatment of p orbitals). Here $n_{0X\xi}$ is the occupation number of the ξth AO of X in a *reference valence state* chosen because of complete knowledge of the corresponding optimum AO's; $F_{X\xi}$ and $K_{X\xi}$ are the well-known Coulomb and exchange operators. One-center integrals appear when the eigenvalues of (21) are evaluated; these one-center integrals and the scaling factor s_X are determined empirically from spectroscopic data using equations derived in part from the virial theorem. These give for a carbon sp^3 state a $Z^* \approx 3.28$, which is different from the crude methane estimate given above. This indicates the extent to which use of a model based exclusively on free-atom orbitals may differ from a reference model where MAO's are, by brute force, adjusted to a molecular situation. In other words, although they recognize the importance of defining a one-electron Hamiltonian for choosing MAO's, Matheu and Ellison (1972) still accept the idea that free-atom properties must be used to evaluate their characteristics, as did Simon and Pujol (1969) in their earlier more direct treatment of shielding effects in valence states.

We have thus formulated the MAO problem and indicated the lines of development which seem to be taking shape. We conclude this section with a last remark, which completes the general picture. The fact that one wants a model with a reference Hamiltonian associated to it implies the very general requirement that both a complete set of MAO's $|\chi_\mu\rangle$ and a complete set of orbital energies ε_μ be available. This can be most easily done by defining a suitable Hamiltonian in the customary form. But it need not be so. As a matter of fact, one can use the well-known expression

$$\mathscr{H}_{\text{model}} = \sum_{\mu} \varepsilon_{\mu} |\chi_{\mu}\rangle\langle\chi_{\mu}|$$

if orbital energies and orbitals are all known; one can use the pseudo-potential idea (Gombas, 1967) if the higher energy part of the basis is known or given by a Hamiltonian having a well-specified form. Indeed, the pseudo-potential idea may be used to find all MAO's, much as it was used to justify Slater's rules (Gombas and Gaspar, 1952).

As regards the "effects" which can be introduced in order to analyze and interpret the transition from FAO's to MAO's, an important step has been made by Somorjai (1968, 1969), who has proposed that the AO's should be treated as transforms of the type

$$\chi(\mathbf{r}) = \int_D S(t)\phi(t, \mathbf{r})\, dt \tag{23}$$

where t is a parameter, D its range of variation, $\varphi(t, \mathbf{r})$ a function of the electron position vector $\mathbf{r}$ which coincides with a free-atom orbital when t takes some value t_0, $S(t)$ a "shape function" which becomes the Dirac δ-function when $\phi(t_0, \mathbf{r}) = \chi(\mathbf{r})$.

VII. The MO–LCMVAO Model

A nonintuitive choice of the MAO basis is the only serious open problem of the whole model treatment of a molecule that may be called "Molecular Orbitals by Linear Combinations of Modified Valence Atomic Orbitals," MO–LCMVAO. The other steps can be easily formulated and provide a consistent scheme which can be improved but is already well defined and easy to connect to higher class computations.

The first question the model must answer is: How is a molecule divided into two-center and many-center bonds (including π systems)? The answer can be given in two steps. First, one has to decide *where* the bonds are; then one must adjust the basis to that kind of description. It is well known that localization can be realized in an infinite number of ways; this is what led some to claim that the bond concept is an arbitrary one. One can eliminate the arbitrariness by introducing hybridization according to the MO version of Pauling's (1931) idea (Del Re, 1963): Combine the pure MAO's for each atom separately so as to obtain, if possible, orthogonal *hybrid* orbitals having as small an overlap as possible with all the other hybrid orbitals except one; the pairs thus obtained will define the bonds. If no matter how hybridization is realized more than two hybrids have overlaps of the same order of magnitude, we shall say that we have a many-center bond.

The procedure to realize the scheme outlined above is as follows: (a) Compute the overlap matrix S for the pure basis and arrange it in blocks $\mathbf{S}^{(XY)}$ associated with pairs of atoms, (b) diagonalize by means of unitary matrices $\mathbf{W}^{(XY)}$ each of the $\mathbf{S}^{(XY)}$ blocks via the symmetric matrix $\mathbf{S}_{XY}^{\dagger}\mathbf{S}_{XY}$:

$$\mathbf{W}^{(XY)\dagger}\,\mathbf{S}^{(XY)\dagger}\,\mathbf{S}^{(XY)}\,\mathbf{W}^{(XY)} = \text{diag}, \tag{23a}$$

$$\mathbf{W}^{(YX)\dagger}\,\mathbf{S}^{(XY)}\,\mathbf{W}^{(XY)} = \text{diag}; \tag{23b}$$

(c) for each atom X of valence v select the v highest eigenvalues of Eq. (23a) allowing Y to be any atom in the molecule (this is the step deciding where the bonds are) and call $\mathbf{W}_{\xi(Y)}$ the columns of the $\mathbf{W}^{(XY)}$ matrices (X fixed) which correspond to those highest eigenvalues, which are denoted by $\lambda_{\xi(Y)}^{2}$; (d) construct v *orthogonal* columns $U_{\xi(Y)}^{(X)}$ using the condition that for equal $\lambda_{\xi(Y)}^{2}$ they should be as close as possible to the corresponding $\mathbf{W}_{\xi(Y)}$'s, and the closer the higher the corresponding λ^2 value; and (e) build a block diagonal unitary transformation $\mathbf{U}$ where each block $\mathbf{U}^{(X)}$ is obtained from the set of v columns $\mathbf{U}_{\xi(Y)}^{(X)}$ just obtained (Del Re, 1963).

If the orthogonality restraint were not there, the above procedure would give a unitary transformation of the pure MAO basis which mixes with one another only the MAO's of the same atom (hybridization) and makes most of the elements of the new overlap matrix as small as is possible with such a kind of transformation. Orthogonalization reduces the quality of the results, but usually does give good results (Veillard and Del Re, 1964): In ordinary saturated molecules it does happen that overlap corresponding to bonds of the chemical formula are large.

(One exception is given by hydrogen–hydrogen overlap in methyl and methylene groups, and this may indicate either that the Slater AO's used are not good approximations to correct MAO's or that CH bonds should not be treated separately.)

The procedure outlined above has the quality of not requiring knowledge of where bonds are; indeed, an orbital-following procedure can be based on it so as to obtain computed geometries. That work has given excellent results for hydrocarbons (Rastelli *et al.*, 1972). However, difficulties with lone pairs appear in heteroatomic molecules, which is not surprising because the overlap criterion conceals an approximation on the Hamiltonian (Del Re, 1963). Ways to overcome that difficulty have been suggested and amount to assigning a "weight" λ_{lp}^2 to lone pair orbitals in step (e) above (Del Re *et al.*, 1966). However, a more elegant procedure can be suggested, based on a version of the above idea originally elaborated for *a posteriori* analysis of *ab initio* calculations (McWeeny and Del Re, 1968). First, one builds an approximate one-electron Hamiltonian **H** on the pure MAO basis, according, say, to the EH method recipes (Hoffman, 1963; Rastelli *et al.*, 1972). Then one constructs a "localized" density matrix **P** corresponding to the bonds singled out according to the procedure outlined above (single elements equal to 2 for lone pairs, 2 × 2 blocks with elements equal to 1 for two-center bonds, etc.). Finally, one introduces an unknown hybridization matrix **U** consisting of unitary blocks which do not combine orbitals of different atoms, and solves the equation

$$\text{Trace } \mathbf{P}\mathbf{U}^{\dagger}\mathbf{H}\mathbf{U} = \max \tag{24}$$

(blocks of **P** correspond to lone pairs, bonds, and π systems). This gives a basis hybridized with explicit account of lone pairs. The solution of Eq. (24) is a straightforward business, as is shown in McWeeny and Del Re (1968).

The basis obtained is entirely equivalent to the original MAO basis. The use of approximations as that involved in using overlap and an EH effective Hamiltonian has no influence on the quality of the basis and can only affect the meaningfulness of the picture presented by the hybrids. The next step in the development of the model consists in deciding what effective molecular Hamiltonian we want to use for our idealized (or simplified) picture of a molecule and to show how it fits in a formalism as rigorous but complete.

The essence of the traditional MO–LCAO method involves use of an effective Hückel-like one-electron Hamiltonian, and use of two-center two-AO bond orbitals for typical σ bonds. In order to obtain a formula where an effective Hamiltonian of that kind appears explicitly consider a one-determinant approximation and then use the general formalism of the Appendix. In order to avoid complications, and to keep in complete line with the current chemical point of view, one requires that the overlap matrix over the final hybrid MAO basis have the same block structure as **P** in Eq. (24). To insure this, several procedures are available (Del Re, 1973). Practical applications (Lami and Del Re, 1974) suggest that the most satisfactory one is based on symmetric orthogonalization (Löwdin

1947, 1950, 1956). The matrix $\mathbf{S}$ over the hybrid basis $|\mathbf{h}\rangle$ obtained according to the procedures just described is divided into two parts $\mathbf{S}_0$ and $\mathbf{S}'$, corresponding to the nonvanishing blocks of $\mathbf{P}$ and to the remaining elements, respectively. Then $\mathbf{S}_0$ is brought to a diagonal form $\mathbf{s}_0$ by a block-diagonal transformation $\mathbf{T}$:

$$\mathbf{T}^{\dagger}\mathbf{S}_0\mathbf{T} = \mathbf{s}_0, \qquad \mathbf{T}^{\dagger}\mathbf{T} = \mathbf{I} \tag{25}$$

Next the new basis $|\mathbf{h}\rangle\mathbf{T}$ thus defined is partially othogonalized to give a final hybrid basis $|\mathbf{h}_f\rangle$ according to the expression

$$|\mathbf{h}_f\rangle = |\mathbf{h}\rangle\,\mathbf{T}\,\mathbf{s}_0^{-1/2}(\mathbf{I} + \Delta\mathbf{S})^{-1/2}\,\mathbf{s}_0^{-1/2}, \qquad \text{with} \quad \Delta\mathbf{S} = \mathbf{s}_0^{-1/2}\mathbf{T}^{\dagger}\mathbf{S}'\mathbf{T}\mathbf{s}_0^{-1/2}, \tag{26}$$

which can be easily computed because $\mathbf{s}_0$ is diagonal and $\Delta\mathbf{S}$ is small (thus allowing use of a series expansion). Of course, the whole procedure can be summarized by the expression

$$|\mathbf{h}_f\rangle = |\chi\rangle\,\mathbf{U}\mathbf{S}^{-1/2}\,\mathbf{S}_0^{1/2} \tag{27}$$

where $\mathbf{U}$ is the hybridization matrix and $|\chi\rangle$ is the initial pure MAO basis.

The new basis $|\mathbf{h}_f\rangle$ is thus adjusted to the molecular situation by a slight delocalization in addition to hybridization. Instead of delocalizing, one can insure the same result by radial distortion according to Rastelli and Del Re (1969); at any rate we have again a basis equivalent to the original one, with rigorous block factorization of the overlap matrix. This means that molecular orbitals obtained from subsets of the basis corresponding to different blocks of $\mathbf{S}$ are automatically orthogonal to one another.

Consider now a determinant Ψ built over the groundstate occupation scheme for the various types of molecular orbitals that can be obtained from $\mathbf{S}$; for instance, in pyridine, one MAO for the lone pair, eleven two-center two-MAO bond orbitals coming each from one of the eleven blocks corresponding to the sigma bonds, three six-center bonds for the π orbitals. If the molecular orbitals coming from the same block (like the π orbitals of pyridine) are orthogonal to one another, then the usual formula for the total energy can be used

$$(\Psi|\mathscr{H}|\Psi) = \sum_{A,B}^{A\neq B}\frac{Z_A Z_B}{R_{AB}} + 2\sum_k \mathbf{C}_k^{\dagger}\mathbf{H}_{\text{core}}^{(k)}\mathbf{C}_k + \sum_{k,l}[2(kk|ll) - (kl|kl)] \tag{28}$$

where k and l refer to molecular orbitals occupied twice in Ψ, $\mathbf{C}_k$ is the column vector of the coefficients of its expansion in the basis $|\mathbf{h}_f\rangle$, $\mathbf{H}_{\text{core}}^{(K)}$ is the block of $\mathbf{H}_{\text{core}}$ associated with the block of $\mathbf{S}'$ to which the kth MO belongs, $(kk|ll)$ and $(kl|kl)$ are the Coulomb and exchange integrals associated with the kth and the lth MO, and R_{AB} is the distance between the atoms A and B, whose effective nuclear charges are Z_A, Z_B.

It is possible at this stage to express the various integrals as sums of simple expressions plus corrections. More specifically, for nuclear attraction integrals,

$$(X\xi|A|Y\eta) = \tfrac{1}{2}S_{X\xi Y\eta}(F_{XA} + F_{YA}) + \Delta(X\xi|A|Y\eta); \tag{29}$$

and for two-electron integrals,

$$\begin{aligned}(X\xi\, Y\eta|Z\zeta Uv) = {} & \tfrac{1}{4}S_{X\xi Y\eta}S_{Z\zeta Uv}(F_{XZ} + F_{XU} + F_{YZ} + F_{XU}) \\ & + \Delta(X\xi\, Y\eta|Z\zeta Uv);\end{aligned} \tag{30}$$

where, as before, $X\xi$ denotes the ξth hybrid MAO of X, $Y\eta$ denotes the ηth hybrid MAO of Y, etc., and F_{AB} is $1/R_{AB}$ unless $A \equiv B$; in the latter case it takes a suitable value.

These partitions can be used to complete the construction of a simple MO–LCAO model within the frame of a correct energy expression.

First of all, consider a single AO group K, namely a group of AO's which are associated with one of the nonvanishing blocks of $\mathbf{S}$, viz. lone pairs, two-center bonds, many-center bonds, delocalized π systems. Denote by X, Y, X', ... atoms which are centers of K, by A, B, A', ... atoms *not* belonging to K. Then one can divide $\mathbf{H}_{\text{core}}^{(K)}$ into three parts and write

$$\begin{aligned}2\sum \mathbf{C}_k{}^\dagger \mathbf{H}_{\text{core}}^{(K)} \mathbf{C}_k = {} & 2\sum_k \sum_{X,\xi} \sum_{Y,\eta} c^*_{X\xi k} c_{Y\eta k}\left(X\xi \left| \mathscr{T} - \sum_{X'} \frac{Z_{X'}}{r_{X'}} \right| Y\eta\right) \\ & - \sum_X \sum_A \frac{Z_A Q_X}{R_{AX}} + 2\sum_k \sum_{X,\xi} \sum_{Y,\eta} \sum_A \Delta(X\xi|A|Y\eta)c^*_{X\xi k} c_{Y\eta k},\end{aligned} \tag{31}$$

where $\mathscr{T}$ is the one-electron kinetic-energy operator, the partition (29) has been applied to the "nonbonded" part of $\mathbf{H}_{\text{core}}$, and Q_X is the total gross atomic population of X determined by summation over all k's of (19) and multiplication by 2. Using Eq. (30) one also obtains

$$(kl|mn) = \sum_X \sum_{X'} Q_X^{(kl)} Q_{X'}^{(mn)} \frac{1}{R_{XX'}} + \Delta(kl|mn) \tag{32}$$

where

$$Q_X^{(kl)} = \operatorname{Re} \sum_{\xi, Y, \eta} c_{X\xi k}^* c_{Y\eta l} S_{X\xi Y\eta} . \tag{33}$$

Equation (32) can be used to obtain an explicit separation into groups of the one-electron part of the energy expression (σ–π separation, bond–bond separation). To do this one can partition the Coulomb part of the two-electron term of Eq. (28) into contributions belonging and not belonging to a given K group, and apply Eq. (32) to the part which corresponds to "nonbonded atoms." Noting that if A is not linked to X it does not belong to K, and $Q_A^{(kk)} = 0$, one can write

$$\begin{aligned}
&\sum_{k,l} [2(kk|ll) - (kk|kk)] \\
&\quad = \sum_k \sum_{X,\xi} \sum_{Y,\eta} c_{X\xi k}^* c_{Y\eta k} \left(X\xi \,\middle|\, \sum_{X'\xi'} \sum_{Y'\eta'} \left[\left(2 \sum_l c_{X'\xi' l}^* c_{Y'\eta' l} - c_{X'\xi' k}^* c_{Y'\eta' k} \right) \right.\right. \\
&\qquad \times \left(X'\xi' \left| \frac{1}{r} \right| Y'\eta' \right) | Y\eta) + \sum_{k,l} \sum_{X,A} 2 Q_X^{(kk)} Q_A^{(ll)} \frac{1}{R_{AX}} + \text{correction},
\end{aligned} \tag{34}$$

where X' is an atom linked to X, A is an atom *not* linked to X. In accordance with the partitioning already introduced, the orbital-product potential can be written in the form

$$\left(X'\xi' \left| \frac{1}{r} \right| Y'\eta' \right) = \frac{1}{2} S_{X'\xi' Y'\eta'} \left(\frac{1}{r_{X'}} - \frac{1}{r_{Y'}} \right) + \Delta(X'\xi' | Y'\eta'). \tag{35}$$

The following identity, clearly related to a well-known expression of the total energy as a sum of energies plus two-electron terms (McWeeny and Sutcliffe, 1969), also holds:

$$\begin{aligned}
&\left(X\xi \left| \sum_{X'\xi'} \sum_{Y'\eta'} \left[2 \sum_l c_{X'\xi' l}^* c_{Y'\eta' l} - c_{X'\xi' k}^* c_{Y'\eta' k} \right] S_{X'\xi' Y'\eta'} \frac{1}{2} \left(\frac{1}{r_{X'}} + \frac{1}{r_{Y'}} \right) \right| Y\eta \right) \\
&\quad = 2 \left(X\xi \left| \sum_{X'} (Q_{X'} - Q_{X'}^{(K)}) \frac{1}{r_{X'}} \right| Y\eta \right) - \frac{1}{2} \sum_{X'} (Q_{X'} - Q_{X'}^{(K)}) S_{X\xi Y\eta} (F_{XX'} + F_{YX'}) \\
&\qquad - \sum_{X'} (Q_{X'} - Q_{X'}^{(K)}) \, \Delta(X\xi | X' | Y\eta),
\end{aligned} \tag{36}$$

where, as before, $Q_X = 2 \sum_l Q_X^{(ll)}$ and $Q_X^{(K)}$ is an average value of $Q_X^{(kk)}$ within the group K. Note that the rearrangement shown is intended to introduce an effective shielding potential of the type (16). The second sum will give a term which is not included in the effective one-electron Hamiltonian. Insertion of expression (36) in (34), of (34) and (31) in (28), gives

$$
\begin{aligned}
(\Psi|\mathscr{H}|\Psi) = {} & \sum_X \left(\frac{1}{2} Q_X{}^2 - \sum_k Q_X^{(kk)2}\right) F_{XX} + \frac{1}{2} \sum_{A,B}^{A \neq B \text{ (n.b.a.)}} \frac{q_A q_B}{R_{AB}} \\
& + \frac{1}{2} \sum_{X,Y}^{X \neq Y \text{ (b.a.)}} (Z_X Z_Y - Q_X Q_Y) F_{XY} \\
& + 2 \sum_k \mathbf{C}_k{}^\dagger \mathbf{H}_{\text{model}}^{(K)} \mathbf{C}_k + \Delta(\Psi|\mathscr{H}|\Psi).
\end{aligned}
\tag{37}
$$

In this expression note the following: (a) The first summation is a self-energy term. (b) The electrostatic repulsion term between nonbonded atoms (sum denoted by "n.b.a.") corresponds to "net" atomic charges

$$q_A = Z_A - Q_A. \tag{38}$$

(c) The major core-repulsion term between bonded atoms ("b.a.") is an electrostatic repulsion *without* cross terms and with a *minus* sign for the gross population product. (d) The matrices $\mathbf{H}_{\text{model}}^{(K)}$ are matrices associated each with the orbital group K corresponding to a particular block K of the overlap matrix over the basis $|\mathbf{h}_f\rangle$; they correspond to effective Hamiltonians

$$\mathscr{H}_{\text{model}}^{(K)} = \mathscr{T} - \sum_{X' \in K} \frac{Z_{X'} - (Q_{X'} - Q_{X'}^{(K)})}{r_{X'}}. \tag{39}$$

These are of the SCF type but become Hückel-type Hamiltonians if the Q's are replaced by standard values, e.g., uniform distributions. (e) The correction to the expression (37) includes all the terms which must be introduced to restore the correct energy value; they fall into general categories, mainly corresponding to neglect of extension and of directionality of AO's (in addition to the Mulliken approximation) in Eqs. (29), (30), etc., and to exchange. These corrections can be intended as "effects" added to some basic electrostatic scheme. They may be very important, but this does not mean that they should not be treated as separate contributions.

The outline given above can be completed by reference to Del Re and Parr (1963), Blyholder and Coulson (1968), Del Re (1972), and to the Appendix. The essential point is that we have rearranged the energy formula so as to sort out model Hamiltonians of the type (39) and effective core-repulsion terms to go with them. This means that we can now associate separate treatments with the single groups of orbitals separated by the procedure leading to the basis $|\mathbf{h}_f\rangle$ and study the corrections in a separate step. [These corrections, of course, include correlation in the form of configuration interaction, as is shown explicitly for π systems in Del Re and Parr (1963).] For instance, we can study a π system according to the Hückel method, or a σ system according to a localized picture with inductive effects (Del Re, 1972, 1973); the parameter problem may be handled again by analyzing the Hamiltonians (39) and possibly separating from them parts which do not correspond to the parameters chosen.

To sum up, Eq. (37) suggests a partitioning of the total energy expression that, with proper care, gives a separation of effects corresponding to bonds and nonbonded interactions, to distortions and hybridization of the basis, to exchange, etc. This makes it possible to apply with confidence the standard technique of physics, which consists in proposing a first interpretation of phenomena on the basis of simplified pictures, following it by inclusion of additional corrections as the need may be.

The essential point is that such a process should not be entirely arbitrary. In the scheme just discussed it is not. For instance, consider a treatment of conformations. If π interactions between formally unlinked atoms are included in the bonded interactions, the steps are:

(1) Reduce the basis to a minimal standard basis.

(2) Decide where and what the bonds are on the basis of overlap.

(3) Determine the optimum hybrids on an approximate form of effective one-electron Hamiltonian.

(4) Block factorize completely the overlap matrix, so as to have different MO group.

(5) Express the total energy as a sum of contributions coming from the individual MO groups corresponding to the different blocks.

(6) Distinguish between intragroup and intergroup contributions to the energy (the latter will be nonbonded interaction energies).

The above procedure is evident when the MO groups in question consist of a bonding and an antibonding bond orbital or reduce to an n center orbital. It is also valid for π systems.

Also a definition of theoretical bond energies becomes possible with a model treatment extracted from Eq. (37) (Del Re, 1972, 1973), as is shown in the next section.

The treatment just presented leads to the definition of an MO model entirely consistent with the traditional interpretation of the chemical formula of a molecule and in particular with the separation into different subsystems (bonds, conjugated π systems, etc.); adoption of that model without further discussion involves neglect of specific terms, which are well defined in the energy expression. It is legitimate to neglect nuclear motions in describing a molecular equilibrium state.

It may be argued that this interpretation of models places some restrictions on parametrization; this is so, and indeed it is well known, although it is not much taken into account in practice owing to the absence of clear-cut analyses.

Once a model has been defined and adopted, further work is done by replacing the actual molecule by the fictitions physical system whose Hamiltonian is precisely the model Hamiltonian. This gives complete freedom as regards parametrization as far as that model is concerned; indeed parametrization is but another way of computing matrix elements. Where the restrictions come into play is comparison with experiment. Just imagine that we try to adjust the parameters of a Hückel scheme to quantities involving exchange as happens with many spectroscopic adjustments; this is wrong, because the model does not admit such an effect as exchange. Similar less evident restrictions exist or may exist for several experimental quantities; this is why the analysis sketched in this section is important and should be pushed further as the need for simplified models increases, even just in connection with *a posteriori* analyses of *ab initio* computations.

In conclusion, the traditional schemes of the traditional theory of chemistry and of elementary quantum chemistry involving such concepts as chemical atoms, bonds, hybridization, inductive effects, σ–π separation, conjugation, etc., are acceptable insofar as they are features of well defined model systems. Indeed, it is legitimate and advisable to introduce models (in particular the MO–LCAO model) to simplify the task of analyzing chemical facts; what is not legitimate is to forget that a model cannot be anything, but must be defined so as to permit precise assessment of what it can and what it cannot simulate.

VIII. Old and New Concepts and an Outlook for the Future

Papers concerning various aspects of the preceding discussion are scattered sparsely in the literature. We mention again Del Re and Parr (1963) and Blyholder and Coulson (1968) as examples; in the former, the Hückel model for a π system is extracted from a general CI formulation; in the latter the extended Hückel method (EHM) approximations are shown to correspond to a scheme which, upon close examination, appears to be a special version of Del Re and Parr (1963) and is completely consistent with the model concept developed in last section.

The fact that the EHM—or a simplified form of it—thus acquires the dignity of a model may surprise some in view of the serious failures of that method. Actually, there is nothing contradictory in this; as has been mentioned, one should not demand that a model give what it cannot give, say conformational energies via the bond integrals. In fact, that simple scheme can give quite good predictions on conformations provided it is based on variation via the Coulomb integrals of the diagonal elements of the effective Hamiltonian rather than of its off-diagonal elements (Cambron-Brüderlein and Sandorfy, 1966).

All semiempirical methods for the study of molecular electronic wavefunctions can be treated by defining a convenient model along the lines indicated above. However, we insist that parametrization should not be confused with choice of a model.

To illustrate in detail how a model can be used to generate quantities associated with current concepts of chemistry we go back to the localized MO–LCAO model illustrated in the preceding section and consider two concepts: inductive effects and bond energies.

The experimental evidence concerning inductive effects is mainly derived from the relative reactivities of organic compounds (Ingold, 1953). Now, an analysis of those data is a very delicate problem, for it involves the theory of chemical reactions. Nevertheless, one can accept the current interpretation of inductive effects as corresponding to differences in charge transfers. The quest for a theoretical counterpart of them then requires: (1) a localized–bond description, because the notion of inductive effect is based on a starting model where bonds are considered as strictly independent building blocks of a molecular system, and (2) a definition of the theoretical quantities that are characteristic of atoms and depend on the presence of several bonds ending at the same atom.

As is well known, total net charges of atoms are good representatives of inductive effects; they can be adjusted to reproduce dipole moments and other quantities of the same type (Del Re, 1958). However, the considerations of the preceding section show that the analysis must go further back in the theory, toward an effective Hamiltonian. In fact, net charges are not necessarily an intermediate step in a theoretical chain, but one of the quantities depending on the entity that really represents inductive effects, namely the effective Hamiltonian.

Inspection of Eq. (39) shows that gross atomic populations must be responsible for inductive effects via the matrix elements of the effective Hamiltonian associated with each bond. This means that the individual

Hamiltonian matrices associated with the various bonds are interdependent. A simple way to introduce that interdependence consists in assuming that

$$H_{XX}^{(K)} = H_{XX}^{\circ(K)} + \sum_{Y \neq X}{}^{(\mathrm{b.a.})} \gamma_{XY} H_{YY}^{(K')}, \tag{40}$$

where we consider only two-center bonds, and therefore denote matrix elements by the atoms and by the bond to which they correspond. The coefficients γ in Eq. (40) are essentially transfer coefficients in the sense of the theory of linear systems.

Equation (40) can be replaced by a much more complicated expression and does not rigorously correspond to Eq. (39). However, at least for small differences between the H_{XX}'s, the charge transfers are linear functions of those differences, and therefore Eq. (40) corresponds to a first-order approximation issued from Eq. (39). In practice direct adjustment of charges to one another and a procedure like Eq. (40) are nearly equivalent (Fliszar, 1972).

The advantage of having an expression like Eq. (40) to represent, so to speak, the source of inductive effects lies in the fact that one can associate the latter with quantities other than charge transfers. For instance, one can try to examine theoretically two problems: deviations from additivity of electric dipole moments and deviations from additivity of bond energies. The former depend on net charges in a very direct way (Del Re, 1967); the latter must be first of all defined in a theoretical sense, because in the empirical scheme they stem from a partitioning of thermochemical data.

An analysis of Eq. (37) gives the following approximate expression for quantities which can be taken as theoretical bond energies:

$$\begin{aligned} E_{XY}^{(K)} \cong 2\varepsilon_{XY}^{(K)} + \frac{Z_X Z_Y}{R_{XY}} - \left(\frac{1}{2m_X} Q_X{}^2 - Q_X^{(K)2}\right) F_{XX} \\ - \left(\frac{1}{2m_Y} Q_Y{}^2 - Q_Y^{(K)2}\right) F_{YY} - (Q_X Q_Y - Q_X^{(K)} Q_Y^{(K)}) F_{XY}, \end{aligned} \tag{41}$$

where $\varepsilon_{XY}^{(K)}$ is the bonding orbital energy, and m_X, m_Y are the numbers of bonds in which X and Y participate (Del Re, 1972, 1973). The dependence of these energies on charge transfers is quite clear, and one sees, among other things, that a bond has a higher energy content the richer it is in electrons. Upon closer inspection one can also see that the orbital energy matters comparatively little in the overall balance of Eq. (41). The conclusion is that, in distributing energy among bonds, the most important molecule-dependent factors are the effective repulsions of atoms (Fliszar *et al.*, 1974).

Other current concepts of chemistry have been more or less directly discussed in the preceding sections, notably hybridization as a result of localization and conjugation as a result of block factorization of the overlap matrix. The model underlying them is the simple MO–LCAO model.

Two questions arise: Is it possible and useful to derive new concepts from the model in question? What about sophisticated *ab initio* computations?

The first question is answered by an example. In the above scheme, the weakest point is the MAO problem, which has not been solved as yet. It may happen that, once rules are found for the change of the pure MAO's of an atom in passing from one molecule to another, one can speak of a special effect associated with them and revealed by some experimental data, say nuclear quadrupole coupling constants and the like (Bonaccorsi *et al.*, 1969).

As regards highly sophisticated computations, they are here to stay, but we insist that they provide raw material to be analyzed, and an analysis requires a model, so that one is back to the model problem. In this connection, some reflections on the future are in order. In an ideal world, the future of "predictive" semiempirical methods should be rather dim; they are engineering interpolation procedures possibly useful, but only justified by success in predictions of one or two properties of a class of molecules (Klopman and O'Leary, 1970; Kutzelnigg *et al.*, 1971, and references therein). Nevertheless, those methods may hide a physical picture; in that case they have a value which is independent of their predicting power, and deeper investigation of them may suggest new viewpoints in the real task of understanding molecules. In this connection we emphasize that the analyses of Sections VI and VII apply to the various methods derived from the Hückel method as well as to schemes like the PPP scheme (Klopman and O'Leary, 1970) and to newer methods, like the one suggested by Slater (1971). The other extreme is highly refined studies on electron correlation (Löwdin, 1969; Herigonte, 1972). Here again, a clear-cut starting model may be invaluable by simplifying the computations in favor of clarity of results. This is the *leitmotiv* of much remarkable work on correlation by the coupled-pair many-electron theory (Paldus and Cizek, 1971).

In addition to extraction of the underlying physical picture from semiempirical methods, development of qualitative rules based on simple models is still necessary to put some order in the overpopulated world of organic chemistry (Parr and Schuveller, 1972).

For the rest, it seems that many fields have been neglected where the present know-how and the highly developed computing facilities would be very useful even in approximate treatments: in particular excited states, nuclear motions, and models of chemical reactions. Here, models to handle the electronic part of the total wavefunction of an isolated molecule are essential in the context of computer simulation of molecular systems (cf., e.g., Malcolme-Lawes, 1972).

Superficial examination of current journals shows that a tendency is already there to shift emphasis to these new topics. The difficulty is that, as has been shown, much remains to be clarified in the fields already greatly exploited; a sudden swing of the pendulum of research could be dangerous, because it might lead to taking for granted a lot of doubtful and possibly unjustified working hypotheses. The present paper has been written with that consideration in mind.

Appendix

The second-quantization formalism makes it possible to use the matrix forms of the one-electron Hamiltonian and of the corresponding Coulomb and exchange operators to write the total Hamiltonian operator. This makes it possible to transfer a partitioning of the energy expression for, say, a groundstate closed-shell wavefunction to the Hamiltonian operator. Of course, it is not necessary for this to use the second-quantization formalism; in fact a proof of this kind was given in the ordinary form in Del Re (1963); but the n representation is definitely more concise and powerful (Longuet-Higgins, 1966).

Consider the general expression for the CI Hamiltonian operator of a molecular problem.

The well-known procedure (Pines, 1963) leads to the formula:

$$\mathscr{H} = \mathscr{R} + a^{\dagger}\overline{\mathbf{H}}^{0}a + \tfrac{1}{2}a^{\dagger}(a^{\dagger}\overline{\mathbf{W}}a)a \tag{42}$$

where $\mathscr{R}$ is the nuclear repulsion multiplier, $a^{\dagger}$ is a row matrix whose elements are the creation operators $a_k{}^{\dagger}$, a is a column matrix whose elements are the annihilation operators a_j, $\overline{\mathbf{H}}^0$ is an ordinary matrix whose elements are the scalar products over the spin orbitals λ_k:

$$(\overline{\mathbf{H}}^{0})_{ij} = \langle \lambda_i | \mathscr{H}^{0} | \lambda_j \rangle, \tag{43}$$

$\mathscr{H}^0$ being the one-electron part of the Hamiltonian (usually $\mathscr{H}_{\text{core}}$), and $\overline{\boldsymbol{W}}$ is a supermatrix whose elements are the two-electron integrals associated with an appropriate two-electron operator $\mathscr{W}(1, 2)$:

$$\begin{aligned}\overline{W}_{ijkl} &= \langle \lambda_i{}^*(1)\lambda_j(1) \,|\, \mathscr{W}(1,2) \,|\, \lambda_k{}^*(2)\lambda_l(2)\rangle \\ &= (ij \,|\, \mathscr{W} \,|\, kl)s(i,j)s(k,l).\end{aligned} \tag{44}$$

In this expression the symbol $s(i, j)$ is zero if the spin parts of λ_i and λ_j are different, one otherwise.

The operator matrix

$$a^\dagger \overline{\mathbf{W}} a = \{\textstyle\sum a_i{}^* \overline{W}_{ijkl} a_j\} \tag{45}$$

is an ordinary matrix whose elements are one-electron operators.

The matrix representation of $\mathscr{H}$ in the Slater determinant or n basis is the usual CI matrix representation which one can write by introducing explicitly the matrix elements of the products of creation and annihilation operators.

If $\boldsymbol{W}$ were zero and $\mathbf{H}$ were a diagonal matrix, the individual n states which form the basis of our representation would already be eigenstates of the system under study; moreover, as they are assumed to be products of spin orbitals conveniently antisymmetrized and possibly spin projected, the result would correspond to what is called an independent-particle model (IPM). More generally, we have a given reference IPM if the spin orbitals used for building the n states are eigenfunctions of some one-electron effective Hamiltonian. This remark is the foundation of our further work. In fact, we intend to show, first of all, that there are many ways of choosing $\overline{\mathbf{H}}^0$ of Eq. (42) because there are many ways of choosing $\mathscr{H}^0$ in definition (43); then we shall discuss semiempirical methods on the basis of the results thus obtained.

We assume first that $\mathscr{H}^0 = \mathscr{H}^{\text{core}}$ and that $\mathscr{W}(1, 2)$ is the electron–electron repulsion operator. To proceed from here, we consider the following identity. The summation $\frac{1}{2}\sum_{i,j,k,l}$ can be divided into

$$-\frac{1}{2}\sum_{j,k} + \sum_{k,l}(\textstyle\sum_j) + \frac{1}{2}\sum_{i,j}^{(i\neq j)}\sum_{k,l}^{(k\neq l)} \tag{46}$$

This means that, if we have

$$\overline{W} = \overline{G} + \overline{F} \qquad [\mathscr{W}(1,2) = \mathscr{G}(1,2) + \mathscr{F}(1,2)]$$

we can always write

$$\frac{1}{2}\sum a_k{}^*a_i{}^*a_j\,a_l\,\bar{F}_{klij} = \frac{1}{2}\sum a_k{}^*a_i{}^*a_j\,a_l\,\bar{G}_{klij}$$
$$-\frac{1}{2}\sum_{j,k}\bar{F}_{jjkk}\,n_j\,n_k + \sum_{k,l}\Bigl(\sum_{j\neq k,l}\bar{F}_{jjkl}\,n_j\Bigr)a_k{}^*a_l$$
$$+\frac{1}{2}\sum_{i,j}^{(i\neq j)}\;\sum_{k,l}^{(k\neq l)}\bar{F}_{ijkl}\,a_k{}^*a_i{}^*a_j\,a_l \tag{47}$$

where account has been taken of the fact that $\bar{F}_{ijkl} = \bar{F}_{klij}$, and that $a_k{}^*a_i{}^*\,a_j\,a_l$ vanishes if $i = k$ or $j = l$.

The above transformation allows one to define, among other schemes, the Hartree–Fock scheme (Pines, 1963). Here we consider another possibility, namely the case when a minimal AO basis is introduced, and special requirements are made as regards the analysis one intends to carry out. In fact, according to Eq. (47), Eq. (42) defines in general an operator of the form

$$\mathscr{H} = \sum \tilde{\bar{H}}_{kl}\,a_k{}^*a_l + \tfrac{1}{2}\sum \tilde{\bar{F}}_{ijkl}\,a_i{}^*a_k{}^*a_l a_j + \tilde{\mathscr{R}}, \tag{48}$$

the bar denoting, as before, matrix taken over the molecular spin orbital basis, the tilde denoting a special choice of the matrices **H** and **F**. In particular, given Slater determinants $|n_L\rangle$, $|n_M\rangle$,

$$(\tilde{\bar{\mathbf{H}}})_{kl} = \Bigl[\bar{\mathbf{H}}^{\circ} + \sum_{j\neq k,l}\bar{\mathbf{F}}_{jj}\,n_j^{(L)}n_j^{(M)}\Bigr]_{kl}, \tag{49}$$

$$\tilde{\bar{F}}_{ijkl} = \bar{G}_{ijkl} + \bar{F}_{ijkl}(1-\delta_{ij})(1-\delta_{kl}), \tag{50}$$

$$\tilde{\mathscr{R}} = \mathscr{R} - \frac{1}{2}\sum_{j,k} n_j^{(L)}n_k^{(M)}n_k^{(L)}n_k^{(M)}\bar{F}_{jjkk}, \tag{51}$$

$$(\bar{F}_{jj})_{kl} = \bar{F}_{jjkl}\,. \tag{52}$$

The occupation number product $n_j^{(L)}n_j^{(M)}$ selects the spin orbitals λ_j which are occupied in both $|n_L\rangle$ and $|n_M\rangle$ (for otherwise it is zero), but does not provide the same selection for all pairs $|n_L\rangle$, $|n_M\rangle$. Therefore, $\tilde{\bar{\mathbf{H}}}$ is not the same matrix for all the matrix elements of the CI Hamiltonian matrix built over the many-electron basis $|n\rangle$.

Suppose we want a single one-electron Hamiltonian representing an idealized but admissible physical system, and thus providing a common, if approximate, (IPM) for all the molecular states. This allows us to give physical significance to concepts like promotion, occupation, etc. In other words, we wish to extract explicitly from the total Hamiltonian Eq. (48) a one-particle Hamiltonian which will be independent of the specific molecular orbitals and of the configurations of Slater determinants which build up the individual matrix elements of the total Hamiltonian: This we shall call a "Hückel" Hamiltonian. For this we consider the matrix **H** in more detail. We have

$$\begin{aligned}(\tilde{\overline{\mathbf{H}}})_{kl} &= (\overline{\mathbf{H}}^{\circ})_{kl} + \sum_{j\neq k,\,l} \langle jj|\mathscr{F}|kl\rangle n_j^{(L)} n_j^{(M)} s(k,\,l)\\ &= \mathbf{C}_k{}^{\dagger}\left[\mathbf{H}^{\circ} + \sum_{j\neq k,\,l}\sum_{\mu,\,\nu} c_{\mu j}^{*} c_{\nu j}\,\mathbf{F}_{\mu\nu}\, n_j^{(L)} n_j^{(M)}\right]\mathbf{C}_l\, s(k,\,l) \qquad (53)\end{aligned}$$

where matrices without a bar refer to the "atomic spin orbital" basis (or to an AO basis $|\chi\rangle$); $c_{\mu j}$ is an element of **C**, and

$$(\mathbf{F}_{\mu\nu})_{\rho\sigma} = (\mu\nu|\rho\sigma) = (\rho\sigma|\mu\nu) \qquad (54)$$

$(\mu\nu|\rho\sigma)$ being the usual two-electron integral over the four AO's indicated in its most general form.

Equation (53) shows that, at variance from $\mathbf{H}^0$, $\tilde{\mathbf{H}}$ is not the representation of a single Hamiltonian operator of the standard type. In fact, it corresponds, in the basis $|\chi\rangle$, to the matrix

$$\tilde{\mathbf{H}} = \mathbf{H}^{\circ} + \sum_{j\neq k,\,l}\sum_{\mu,\,\nu} c_{\mu j}^{*} c_{\nu j}\,\mathbf{F}_{\mu\nu}\, n_j^{(L)} n_j^{(M)} \qquad (55)$$

and this changes not only according to the coefficients of the molecular orbitals appearing in the two vectors $|n_L\rangle$ and $|n_M\rangle$ which gave rise to the matrix element of $\mathscr{H}$ we are analyzing and to the way the electrons are distributed among the various spin orbitals, but according to the pair of spin orbitals (λ_k, λ_l) on which each given element of $\tilde{\mathbf{H}}$ is to be calculated. These are the two points which make the search for a true Hückel-type model difficult especially for σ systems, the more so because such a model will not be quantitatively very satisfactory. A general analysis has been performed by Staemmler and Del Re (1974).

Our problem is to partition $\tilde{\mathbf{H}}$ so as to divide it into two parts, one a "constant" operator, the other a variable perturbation depending on the states over which the matrix elements are taken. To do this, we must replace

$$\sum_{j\neq k,\,l} (\mathbf{P}^{(jj)})_{\mu\nu} = \sum_{j\neq k,\,l} c_{\nu j}^{*} n_j c_{\mu j} \qquad (56)$$

by a constant matrix which may be called $\mathbf{P}_{\mu\nu}^{0}$. As has been mentioned, we require here two different approximations: one on the MO's and one on the occupation numbers. We have to choose $C_{\nu j}^{(0)}$ and $n_j{}^0$; the former are MO coefficients associated with a convenient model, the latter are the occupation numbers of a given reference configuration—say the ground-state configuration. We replace $\mathbf{P}^{(jj)}$ by an average density matrix

$$\mathbf{P}^{(\text{av.},\, jj)} = \frac{1}{n}\,\mathbf{P}^0, \qquad \text{where} \quad P_{\mu\nu}^0 = \sum_j P_{\mu\nu}^{(0,\, jj)} = \sum_j n_j{}^0 C_{\nu j}^{0*} C_{\mu j}^0 . \tag{57}$$

Then we note that the summation in Eq. (56) contains $n - 1$ terms, because either k and l are equal or k refers to a spin orbital λ_k which is empty in $|n_M\rangle$ whereas λ_l is full. Thus, $\mathbf{P}^0$ is a reference density matrix associated with a single configuration on orbitals which correspond to a particular model of the system under study. This model is so chosen as to incorporate certain features and to provide a certain degree of approximation for a group of states of the given system. This means that we can write

$$\sum_{j \neq k,\, l} C_{\mu j}^* C_{\nu j} n_j^{(M)} n_j^{(L)} \mathbf{F}_{\mu\nu} = \frac{n-1}{n}\, P_{\mu\nu}^0 \mathbf{F}_{\mu\nu} + \sum_j (\mathbf{P}^{(jj)} - \mathbf{P}^{(\text{av.}\, jj)})\mathbf{F}_{\mu\nu} \tag{58}$$

Using Eq. (58) we can rewrite Eq. (53) in the form

$$\tilde{\tilde{H}}_{kl} = [\mathbf{C}_k{}^\dagger \mathbf{H}^{\text{eff}} \mathbf{C}_l + \mathbf{C}_k{}^\dagger \mathbf{K} \mathbf{C}_l - \mathbf{C}_k{}^\dagger \{(\mathbf{P}^{(ll)} + \mathbf{P}^{(kk)}) \cdot\cdot\, \boldsymbol{F}\} \mathbf{C}_l] s(k, l) \tag{59}$$

where $A \cdot\cdot\, \boldsymbol{B}$ stands for a matrix whose general element is

$$(\mathbf{A} \cdot\cdot\, \boldsymbol{B})_{\rho\sigma} = \sum_{\mu,\nu} A_{\mu\nu} B_{\mu\nu\rho\sigma} .$$

Comparison of (59) with (58) and (53) indicates that

$$\mathbf{H}^{\text{eff}} = \mathbf{H}^0 + \mathbf{P}^0 \cdot\cdot\, \boldsymbol{F}, \tag{60}$$

$$\mathbf{K}^{(j)} = \{(\mathbf{P}^{(jj)} - \mathbf{P}^{0(jj)}) \cdot\cdot\, \boldsymbol{F}\}, \tag{61}$$

$$\mathbf{K} = \sum_j \mathbf{K}^{(j)} \tag{62}$$

This divides $\tilde{\mathbf{H}}$ into three contributions: (1) $\mathbf{H}^{\text{eff}}$, a matrix which does not depend on the configuration chosen and depends on molecular orbitals which do not change according to the configuration, (2) a "fluctuation" term, which depends on the choice of the configuration and, of course, on the molecular orbitals specific of it, and (3) a term which depends also on the indices of the matrix elements of $\tilde{\mathbf{H}}$ specifically considered.

At this point $\mathbf{H}^{\text{eff}}$ can be taken as the matrix representation of a standard effective Hamiltonian operator which might well be adjusted to reproduce Eq. (39) for the ground state.

A similar analysis can be carried out for $\mathcal{R}$ and $\boldsymbol{W}$ of Eq. (42).

ACKNOWLEDGMENT

This paper was partly written when the author was staying at the Quantum Theory Group of the University of Waterloo, Waterloo, Ontario, Canada. The author thanks Dr. J. Cizek and Dr. J. Paldus for their warm hospitality and stimulating discussions and remarks. The support of C.N.R. (Italy) is also acknowledged.

REFERENCES

ANDRONOV, A. A., VITT, A. A., and KHAIKIN, S. E. (1966). "Theory of Oscillators," Introduction. Pergamon, Oxford.

BERTHIER, G. (1963). *Tetrahedron* **19**, Suppl. 2, 1.

BERTHIER, G., VEILLARD, A., and DEL RE, G. (1964). *Phys. Lett.* **8**, 313.

BERTHIER, G., DEL RE, G., and A. VEILLARD, (1966). *Nuovo Cimento* **44**, 315.

BLYHOLDER, G., and COULSON, C. A. (1968). *Theor. Chim. Acta* **10**, 316.

BONACCORSI, R., SCROCCO, E., and TOMASI, G. (1969). *J. Chem. Phys.* **50**, 2940.

CAMBRON-BRÜDERLEIN, H., and SANDORFY, C. (1966). *Theor. Chim. Acta* **4**, 224.

CLEMENTI, E. (1967). *J. Chem. Phys.* **46**, 3842.

CLEMENTI, E., and POPKIE, H. (1972a). *J. Chem. Phys.* **57**, 4870.

CLEMENTI, E., and POPKIE, H. (1972b). *J. Amer. Chem. Soc.* **94**, 4057.

CLEMENTI, E., and RAIMONDI, D. L. (1963). *J. Chem. Phys.* **38**, 2686.

CLEMENTI, E., and ROUTH, A. (1972). *Int. J. Quantum Chem.* **6**, 525.

COULSON, C. A., and LONGUET-HIGGINS, C. H. (1947a). *Proc. Roy. Soc. Ser. A* **191**

COULSON, C. A., and LONGUET-HIGGINS, C. H. (1947b). *Proc. Roy. Soc., Ser. A* **192**, 16.

COULSON, C. A., and LONGUET-HIGGINS, C. H. (1948). *Proc. Roy. Soc., Ser. A* **193**, 447.

CUSACHS, L. C., and ALDRICH, H. S. (1971). *Chem. Phys. Lett.* **12**, 197.

DEL RE, G. (1958). *J. Chem. Soc. London.* p. 4031.

DEL RE, G. (1963). *Theor. Chim. Acta* **1**, 188.

DEL RE, G. (1964). *In* "Electronic Aspects of Biochemistry" (B. Pullman, ed.), Academic Press, New York, p. 221.

DEL RE, G. (1967). *Int. J. Quantum Chem.* **1**, 293.

DEL RE, G. (1969). *Acta. Phys. Hung.* **27**, 477.

DEL RE, G. (1971). *In* "Aromaticity, Pseudoaromaticity, Antiaromaticity" (E. D. Bergmann and B. Pullman, eds.), p. 74. Isr. Acad. Sci. Hum., Jerusalem.

DEL RE, G. (1972). *Gazz. Chim. Ital.* **102**, 929.

DEL RE., G. (1973). *Int. J. Quantum Chem.* **7**, 193.

DEL RE, G., and PARR, R. G. (1963). *Rev. Mod. Phys.* **35**, 604.

DEL RE, G., ESPOSITO, U., and CARPENTIERI, M. (1966). *Theor. Chim. Acta* **6**, 36.

DEWAR, M. J. S. (1963). *Rev. Mod. Phys.* **35**, 504.

DEWAR, M. J. S., and HASELBACH, H. (1970). *J. Amer. Chem. Soc.* **92**, 3.

ENGLAND, W., and GORDON, H. S. (1971). *J. Amer. Chem. Soc.* **93**, 4649.

FEYNMAN, R. P. (1966). "Lectures on Physics." Addison-Wesley, Reading, Massachusetts.

FLISZAR, S. (1972). *J. Amer. Chem. Soc.* **94**, 1068.

FLISZAR, S., DEL RE, G., and LELJ-GAROLLA, F., (1974). To be published.

GILCHRIST, T. L., and STORR, R. C. (1972). "Organic Reactions and Orbital Symmetry." Cambridge Univ. Press, London and New York.

GOEPPERT-MAYER, M., and SKLAR, A. L. (1938). *J. Chem. Phys.* **5**, 645.

GOMBAS, P. (1967). "Pseudopotentiale," Springer-Verlag, Berlin and New York.
GOMBAS, P., and GASPAR, R. (1952). *Acta Phys. Hung.* **3**, 317.
HALL, G. G. (1954). *Trans. Faraday Soc.* **50**, 773.
HALL, G. G. (1959). *Rep. Progr. Phys.* **22**, 1.
HALL, G. G. (1973). *Chem. Phys. Lett.* **6**, 501.
HERIGONTE, P. V. (1972). *Struct. Bonding (Berlin)* **12**, 1.
HOFFMAN, R. (1963). *J. Chem. Phys.* **39**, 1397.
HÜCKEL, E. (1931). *Z. Phys.* **70**, 204.
INGOLD, C. K. (1953). "Structure and Mechanism in Organic Chemistry." Cornell Univ. Press, Ithaca, New York.
KEMBLE, E. C. (1958). "The Fundamental Principles of Quantum Mechanics." Dover, New York.
KLOPMAN, G., and O'LEARY, B. (1970). *Fortschr. Chem. Forsch.* **15**, 445.
KUTZELNIGG, W., DEL RE, G., and BERTHIER, G. (1971). *Fortschr. Chem. Forsch.* **22**, 86.
LAMI, A., and DEL RE, G. (1974). To be published.
LÖWDIN, P.-O. (1947). *Ark. Mat. Astron. Fys.* **35A**, No. 9.
LÖWDIN, P.-O. (1950). *J. Chem. Phys.* **18**, 365.
LÖWDIN, P. O. (1955). *Phys. Rev.* **97**, 1474, 1490, 1509.
LÖWDIN, P. O. (1956). *Advan. Phys.* **5**, 1.
LÖWDIN, P. O. (1969). *Advan. Chem. Phys.* **14**, 283.
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.
MCLEAN, A. D., and YOSHIMINE, M. (1967). *IBM J. Res. Develop.* **11** Suppl. 12.
MCWEENY, R. (1954). *Proc. Roy. Soc. Ser.* **A 223**, 63.
MCWEENY, R. (1960). *Rev. Mod. Phys.* **32**, 335.
MCWEENY, R., and DEL RE, G. (1968). *Theor. Chim. Acta* **10**, 13.
MCWEENY, R., and SUTCLIFFE, B. T. (1969). "Methods of Molecular Quantum Mechanics," Academic Press, New York.
MALCOLME-LAWES, D. J. (1972). *J. Chem. Soc. Faraday Trans.* **68**, 966.
MALRIEU, B., CLAVERIE, P. and DINER, S. (1969). *Theor. Chim. Acta* **13**, 1.
MATHEU, F. M., and ELLISON, F. O. (1972). *J. Chem. Phys.* **57**, 5183.
MOCCIA, R. (1964). *J. Chem. Phys.* **40**, 2164.
MULLIKEN, R. S. (1932). *Phys. Rev.* **40**, 55.
MULLIKEN, R. S. (1932b). *Phys. Rev.* **41**, 49.
MULLIKEN, R. S. (1933). *Phys. Rev.* **43**, 279.
MULLIKEN, R. S. (1948). *Phys. Rev.* **74**, 736.
MULLIKEN, R. S. (1949). *J. Chim. Phys. Physicochim. Biol.* **46**, 497.
MULLIKEN, R. S. (1951). *J. Chem. Phys.* **19**, 900.
MULLIKEN, R. S. (1955). *J. Chem. Phys.* **23**, 1833, 1841, 2338, 2343.
MULLIKEN, R. S. (1960). *Rev. Mod. Phys.* **32**, 232.
MULLIKEN, R. S. (1966a). *J. Amer. Chem. Soc.* **88**, 1849.
MULLIKEN, R. S. (1966b). *In* "Quantum Theory of Atoms, Molecules, and the Solid State" (P.-O. Löwdin, ed.), p. 231. Academic Press, New York.
NELANDER, B., and DEL RE, G. (1970). *J. Chem. Phys.* **52**, 5225.
O'MALLEY, T. F., and GELTMAN, S. (1965). *Phys. Rev.* *A***137**, 1344.
PALDUS, J. and CIZEK, J. (1971). *Int. J. Quantum Chem.* **5**, 359.
PARR, R. G., and SCHUVELLE, G. W. (1972). *J. Amer. Chem. Soc.* **94**, 8974.
PAULING, L. (1931). *J. Amer. Chem. Soc.* **53**, 1367, 3255.

PAULING, L. (1945). "The Nature of the Chemical Bond." Cornell Univ. Press, Ithaca, New York.

PINES, D. (1963). "Elementary Excitations in Solids." Benjamin, New York.

PULLMAN, B., and PULLMAN, A. (1951). "Les Théories Électroniques de la Chimie Organique." Masson, Paris.

PULLMAN, B., and PULLMAN, A. (1960). *Rev. Mod. Phys.* **32**, 428.

RANSIL, B. J. (1963). *Rev. Mod. Phys.* **32**, 245.

RASTELLI, A., and DEL RE, G. (1969). *Int. J. Quantum Chem.* **3**, 553.

RASTELLI, A., POZZOLI, A. S., and DEL RE, G. (1972). *J. Chem. Soc. Perkin Trans. 2*, p. 1571.

RICHARDS, W. G., WALKER, T. E. H., and HINKLEY, R. K. (1971). "A Bibliography of *ab-initio* Molecular Wavefunctions." Oxford Univ. Press (Clarendon), London and New York.

ROOTHAAN, C. C. J. (1955). *L.M.S.S. Tech. Rep., Univ. Chicago.*

RUEDENBERG, K. (1962). *Rev. Mod. Phys.* **34**, 326.

SALAHUB, D. R., and SANDORFY, C. (1971). *Theor. Chim. Acta* **20**, 227.

SANDORFY, C. (1955). *Can. J. Chem.* **33**, 1338.

SIMON, J. C., and PUJOL, L. (1969). *Theor. Chim. Acta* **21**, 191.

SLATER, J. C. (1930). *Phys. Rev.* **36**, 57.

SLATER, J. C. (1931a). *Phys. Rev.* **37**, 481.

SLATER, J. C. (1931b). *Phys. Rev.* **38**, 1109.

SLATER, J. C. (1963). "Quantum Theory of Molecules and Solids," Vol. 1, McGraw-Hill, New York.

SLATER, J. C. (1971). *Int. J. Quantum Chem.* **55**, 403.

SOMORJAI, R. L. (1968). *Chem. Phys. Lett.* **2**, 399.

SOMORJAI, R. L. (1969). *Chem. Phys. Lett.* **3**, 395.

SMITH, D. W. (1972). *Position Pap. Summer Res. Conf. Theor. Chem. Boulder, Colo.*

STAEMMLER, V., and DEL RE, G. (1974). To be published.

STREITWIESER, A. (1961). "Molecular Orbital Theory for Organic Chemists." Wiley, New York.

VAN VLECK, J. H. (1933). *J. Chem. Phys.* **1**, 177, 219.

VAN VLECK, J. H. (1934). *J. Chem. Phys.* **2**, 20.

VEILLARD, A., and DEL RE, G. (1964). *Theor. Chim. Acta* **2**, 55.

WAHL, A. C. (1966). *Science* **151**, 961.

WOLFSBERG, M., and HELMHOLZ, L. (1952). *J. Chem. Phys.* **20**, 837.

Photoelectron Spectra Showing Relaxation Effects in the Continuum and Electrostatic and Chemical Influences of the Surrounding Atoms

CHRISTIAN KLIXBÜLL JØRGENSEN
Département de Chimie Physique
Université de Genève
Geneva, Switzerland

I. Photoelectron Spectrometry of Gaseous and Solid Samples

The lowest ionization energy I of a gaseous atom or molecule or of a solid has very different, and more well-defined, characteristics than I of the subsequent penultimate orbitals. Thus, the lowest I of a gaseous molecule can be determined (Watanabe, 1957) as the threshold photon energy sustaining ionization and passage of current through the gas, and the lowest I of a solid can be found as the threshold photon energy for Einstein photoemission though such measurements are easier performed on metals than on nonconducting samples. Seen from a theoretical point of view, the lowest I is the lower limit of the *continuum* where the distribution of eigenvalues E of the Schrödinger equation is overall dense, and all the subsequent I values can, at the best, constitute singularities in the continuum.

Since about 1962, the subsequent I values are actually determined as maxima (in the following called signals) in the probability distribution of kinetic energy E_{kin} of the electron ejected with photons of the energy $h\nu$ via the relation

$$I = h\nu - E_{\mathrm{kin}}, \tag{1}$$

and it is customary to plot the corresponding photoelectron spectra with E_{kin} increasing toward the right and I increasing toward the left. It is convenient to use photon sources not needing monochromators emitting a single narrow line with as weak a continuous background as possible. Thus, the resonance line of helium at 21.2 eV and the 2p → 1s emission line of He^+ at 40.8 eV have been used on gaseous samples (Turner *et al.*, 1970). As soft X rays, the 2p → 1s lines of a magnesium (1253.6 eV) or of an aluminum (1486.6 eV) anticathode can be used on both gases (Siegbahn *et al.*, 1969) and solids. Obviously, the photon energy constitutes a higher limit to I measured. The resolution obtained with 21.2 eV photons can be better than 0.01 eV showing vibrational structure in many cases. The photoionization follows the principle of Franck and Condon, the primary process being so rapid as not to allow the internuclear distances to vary. The MO results predicting one-electron eigenvalues below the loosest bound orbital have been magnificently confirmed for diatomic and triatomic molecules (Turner *et al.*, 1970).

Some polyatomic molecules show a complicated vibrational structure producing difficulties for interpretation of the kind well known from ultraviolet spectra (Herzberg, 1966). Thus, the ten different I values (between 9.3 and 25.8 eV) of the fifteen MO's containing the thirty valence electrons of benzene have been determined but the detailed assignment is still discussed (Lindholm *et al.*, 1973; Potts *et al.*, 1973). The deviations from the theorem of Koopmans discussed below are not always exactly the same factor (0.90 to 0.92) multiplying the calculated MO energies. In the special case of methane, the $1a_1$ orbital consisting almost exclusively of carbon 1s has $I = 290.8$ eV, the $2a_1$ orbital $I = 23$ eV, but the three $1t_2$ orbitals (which remain degenerate in the tetrahedral point group T_d of the ground state) produce a complicated structure centered around $I = 14$ eV. A comparison with the isotopically substituted CD_4 shows that this structure is due to the alternative symmetries adapted by the Jahn–Teller unstable CH_4^+. The neon atom shows $I(1s) = 870.2$ eV, $I(2s) = 48.4$ eV, and $I(2p) = 21.6$ eV whereas the ten-electron system H_2O with much lower symmetry (in the point group C_{2v} all three Cartesian axes are nonequivalent) has five distinct "vertical" (i.e. for unchanged internuclear distances) $I = 539.7, 32.2, 18.6, 14.8$, and 12.6 eV.

When energetic photons such as 1486.6 eV X rays are stopped by a gaseous sample, the primary signal (such as $I = 870.2$ eV corresponding to the formation of the highly excited electron configuration $1s2s^22p^6$ of neon) is accompanied, at lower E_{kin} (as would correspond to I values between 900 and 1000 eV) by many weak signals (Siegbahn *et al.*, 1969), due to *shake-up* forming higher excited configurations of Ne^+ such as $1s2s^22p^53p$ and to *shake-off* loosing more than one electron and forming configurations such as $1s2s2p^6$ of Ne^{+2} or $1s2s2p^5$ of Ne^{+3}. These weak signals should not be confused with instances of decreased E_{kin} due to inelastic collisions between the photoelectron and another neon atom or with *Auger electrons* of which E_{kin} does not depend on the original photon energy but represents the energy difference between the original state $1s2s^22p^6$ formed by the photoionization and a subsequently produced electron configuration $1s^22s^22p^4$ of Ne^{+2} loosing the Auger electron. In heavier atoms or molecules, it is sometimes possible to loose several Auger electrons in a cascade of consecutive processes. As known from radioactive transmutations in gaseous molecules, a result of the accumulating positive charge is rapid dissociation to monoatomic ions or oligo-atomic fragments.

Whereas signals due to Auger electrons sometimes are detected in photoelectron spectra of solids, shake-up and shake-off satellites at apparent higher I are not normally conspicuous with the exception of a few elements in definite oxidation states such as copper(II) and lanthanum(III) discussed below. The main reason is the very intense background in photoelectron spectra of solids. Actually, the number of electrons counted per second, when plotted as a function of E_{kin}, forms a staircase structure, each signal being accompanied to the left by an almost constant contribution about half as strong as the signal itself. Most of this background is due to inelastically scattered electrons having indeterminately all kinetic energies below the value given by Eq. (1) though a great amount of electrons with low E_{kin} (say below 100 eV) are also Auger electrons. The fact that the integrated intensity of the background is about 100 times higher than the sum of the signal intensities above the background is not exactly determined by the fact that the average penetration depth (Jørgensen, 1971a) of 1486.6 eV photons in typical samples is 3000 Å, more than a hundred times more than the average escape depth (typically 20 Å) of the electrons not suffering inelastic scattering. Because of the Auger process, each photon produces more than one electron, but only a minor part leaves the surface of the sample.

Metallic samples in electric contact with the apparatus achieve the same Fermi level frequently said to be 4.6 eV. Consequently, the instrument records I^* as the difference between 1482 eV (in the case of 1486.6 eV photons) and E_{kin}. It is frequently argued that $I^* = 284.0$ eV of the carbon 1s signal of graphite or $I^* = 83.8$ eV of the gold $4f_{7/2}$ signal of the metallic element can be used for standardization. However, in the case of nonconducting samples, considerable difficulties are produced by the *charging effects*. During the irradiation with soft X rays, a quasi-stationary positive potential V is obtained by the sample, electrons from the surroundings replacing the ejected photoelectrons. It was suggested (Jørgensen, 1971a) to distribute nonconducting powders on the adhesive side of one-sided Scotch Tape and use the carbon 1s signal of the noncovered hydrocarbon as internal reference with $I(1s) = 290.0$ eV. This is supported by the knowledge (Siegbahn *et al.*, 1969) that carbon atoms in gaseous molecules, when surrounded only by hydrogen atoms and other carbon atoms, have I between 290 and 291 eV (to be compared with 301.8 eV for CF_4 and 297.7 eV for CO_2) and in view of the possibility that I of the solid (polymerized isoprene) is slightly lower than of the gaseous hydrocarbons. However, the polyether on double-sided Scotch Tape or other brands containing polyvinylchloride have higher I(C1s) values. It was suggested (Jørgensen, 1971a) to define a correction C_{st} as the difference between 290 eV and I^* of the hydrocarbon C1s signal. Then, the ionization energy *I relative to vacuo* is taken to be the sum of I^* recorded by the instrument and C_{st}. For comparison with MO calculations and with the data from atomic spectroscopy, it is an obvious advantage to consider I and not I^* which, anyhow, are not defined for nonconductors. Unfortunately, I is overestimated by the expression $I^* + C_{\mathrm{st}}$ because of the quasi-stationary positive potential V. The way out of this difficulty was indicated by Bremser and Linnemann (1971) emphasizing the presence of two carbon 1s signals for typical nonconducting samples on Scotch Tape at a distance frequently between 2 and 5 eV (the latter for certain anhydrous fluorides). Whereas it would have been quite plausible to ascribe the signal at higher I^* to other carbon compounds (such as carbonates and alcohols, including cellulose) it is now realized (Jørgensen and Berthou, 1972a) that in most cases, the distance between the two signals represents V and that corrected ionization energies I' can be defined as the sum of I^* of the individual signals and the difference C'_{st} between 290 eV and the higher I^*(C1s) produced by the Scotch Tape hydrocarbon in intimate contact with the nonconducting sample having the positive potential V. It is known from mixtures of metals such as Pd, Au, CuS and Tl_2O_3 with insulators such as MgF_2, $BaSO_4$, and ThF_4 that the I^* values of the metals are indeed increased 1.5–4 eV corresponding to *V* on the salts (Jørgensen and Berthou, 1972a; Hnatowich *et al.*, 1971). A major argument for the physical significance of the *I'* values is that they are reproducible with an average deviation of 0.2 eV whereas the reproducibility of *I* values is not better than 0.5 eV. Nevertheless, it cannot be argued at present that the absolute ionization energies relative to *vacuo* are known more accurately for nonmetallic solids than $\frac{1}{2}$ to 1 eV, whereas they are known an order of magnitude better for gaseous molecules (Turner *et al.*, 1970).

Measurements with 21.2 eV photons on gaseous molecules (Turner *et al.*, 1970) show areas of vibrational structure of each signal roughly proportional to the number of electrons in the MO ionized, and it is generally easy to make a distinction between two or four electrons having the same *I*. When 40.8 eV photons are used, the MO of the 2p group compounds formed essentially from 2s orbitals in the LCAO approximation give somewhat weak signals. The situation is entirely different with 1486.6 eV photons. Wagner (1972) using a spectrometer Varian IEE-15 like ours determined the intensity of the strongest signal of each element relative to fluorine 1s. In the following, we use the colloquial abbreviation "w" for this unit. It is easy to establish a set of secondary standards of intensity, such as 0.4 w for nitrogen 1s or chlorine $2p_{3/2}$, 2.0 w for sodium 1s, 4.0 w for iodine $3d_{5/2}$, and 6.5 w for caesium $3d_{5/2}$. With exception of thorium and uranium $4f_{7/2}$ running close to 7 w, the latter signal is the strongest known. When comparing with weaker *nlj* signals (Jørgensen and Berthou, 1972a, 1973), we frequently find intensities down to 0.01 w, and in particular, the delocalized MO constituting the major subject of 21.2 eV studies give frequently so weak signals that they are not detected. Thus, hafnium(IV) mandelate $Hf(C_6H_5CHOHCO_2)_4$ has a strong (0.9 w) hafnium $4f_{7/2}$ signal whereas the rather flat region between $I = 9$ and 15 eV gives higher limits to the signals of the 232 valence electrons. It must, in all fairness, be realized that they are distributed over a broad interval with oxygen 2s close to $I = 30$ eV. However, the phenyl "π" electrons and oxygen lone-pair electrons having *I* around 10 eV have at most 0.006 w each. This may be compared with cubic MgO (free of *l*-mixing in the orbitals) where each oxygen 2s electron has 0.01 w and each oxygen 2p electron (not the shell) 1.5 mw. Most signals have roughly the same intensities using 1253.6 eV photons, though a few are half or twice as strong.

The theoretical description of these intensities is fairly complicated though a major component is the probability (intrinsically proportional to the number $(2j+1)$ of electrons) for the X ray to ionize a definite *nlj* value. It is known that high *l* values have a particular high probability of ionization. The intensities in wagner units (Jørgensen and Berthou, 1973) are approximately proportional to $\langle r^{-2} \rangle$ values for Hartree–Fock functions kindly supplied by Dr. R. E. Watson. From the point of view of chemical analysis, it must be realized that the signal intensities are not reproducible better than 20–30%. Hence, the relative concentrations of different elements are, at best, semiquantitative information, but it is interesting that the results apply to the outermost 30 Å of the sample, and also that chemically highly nonequivalent atoms of the same element can be recognized.

The half-sided *half-width* δ of the signals having the shape of a Gaussian error curve is reproducible within 0.1 eV. Certain signals have a standard δ only dependent on the line width of the X rays and on the apparatus, such as the 4f signals in elements heavier than hafnium (Jørgensen, 1972). Thus, the analyzer potential 100 V and 1486.6 eV photons produce $2\delta = 1.9$ eV whereas the optimal conditions (having certain other disadvantages) of analyzer potential 36 V and a high-energy source of 1253.6 eV yield $2\delta = 1.3$ eV. Beyond the possibility of a dispersion of slightly nonequivalent atoms of the same element, the main reason for broadened signals is Heisenberg's uncertainty principle when the ionized state exists for a very short time, emitting X rays or Auger electrons. There are good reasons to believe that the observed δ is the square root of the sum of various squared contributions. Since the Heisenberg contribution to δ is 2.3 eV divided by the half-life in the unit 10^{-16} sec, such a short half-life becomes important for our purposes below 10^{-15} sec where it can increase δ from 0.95 to 0.98 eV or from 0.7 to 0.74 eV. Fadley and Shirley (1970a) pointed out that the $4d_{5/2}$ and $4d_{3/2}$ signals of lutetium(III) in LuF_3 have $\delta = 2.1$ eV corresponding to a Heisenberg contribution 1.9 eV or a half-life $1.2.10^{-16}$ sec. The reason for the short half-life is the transitions first described by Coster and Kronig where a 4f electron rapidly fills the 4d vacancy. The Lu(IV) $4f^{13}$ formed picks up an electron from the surroundings relatively more slowly, as seen from the moderate δ of the Lu 4f signal. Interestingly enough, δ of the $4d_{5/2}$ signal remains almost constant 2.2 eV in all the elements (Jørgensen and Berthou, 1972a) between hafnium and bismuth. The radiative half-life of systems ionized in inner shells is of concern to X-ray spectrometry and is theoretically studied (Bambynek *et al.*, 1972). The broadening due to rapid $n(l+1) \rightarrow nl$ transition can also be seen in 2s, 3s, and 4s signals frequently having δ between 2 and 3 eV and in the most extreme case known, δ of the 4p signal increasing from 2 eV in rhodium(III) to 5 eV in cadmium(II) and decreasing again to 1.3 eV in barium(II). The corresponding half-life 5.10^{-17} sec of the 4p vacancy in cadmium is the shortest we have detected (Jørgensen and Berthou, 1972a). When short half-life of the ionized state is the predominant source of broadening, the signals have Lorentzian shape (inversely proportional to $(I - I_0)^2 + \delta^2$).

A third origin of broadened signals which should not be neglected is the influence of *Franck and Condon's principle*. When the primary excitation lasts for less than 10^{-13} sec (exactly like visible and ultraviolet spectra) the nuclei do not have sufficient time to move away from their positions in the electronic groundstate. Even when the vibrational structure of a photoelectron signal of a gaseous molecule is not resolved, δ may still be above 0.5 eV though other signals belonging to nonbonding orbitals may be far narrower (Turner *et al.*, 1970). When we observe $\delta = 1.3$ eV for lithium 1s and a marginal broadening to $\delta = 1.1$ eV for beryllium 1s, it is probably connected with the Franck–Condon principle, the internuclear equilibrium distances being shorter in the 1s systems Li(II) and Be(III) with decreased ionic radii. It is also generally true (Jørgensen and Berthou, 1972a) that anions tend to have broader signals than isoelectronic alkaline ions, probably because fluoride changes the ionic radius more than sodium as a consequence of the ionization to $1s2s^22p^6$, and a certain influence of the chemical bonding on δ of the 1s signals of several nitrogen and oxygen compounds may be explained along the same line, the Franck–Condon contribution of δ in some cases being close to 1 eV.

One might have expected the outermost atoms in a given compound to have changed ionization energies because of differing Madelung potentials and conditions of chemical bonding. Since the order of magnitude of the Madelung potential in uni-univalent salts is 10 eV, it is not trivial that one at most finds a residual contribution to δ around 0.3 eV under the best conditions of resolution, when compared with genuine metals such as gold. It is true, however, that the adsorption of monomolecular layer of various molecules on gold seems to increase $I^*(4f)$ about $\frac{1}{2}$ eV.

II. Orbital Ionization Energies and Anti-Koopmans Relaxation of the Other Orbitals

In the gaseous molecules (Turner *et al.*, 1970) a distinction is made between "adiabatic" ionization energies allowing the internuclear distances to vary (as done when evaluating thermodynamical quantities such as standard oxidation potentials) and "vertical" ionization energies keeping the internuclear distances invariant. In this review, all the photoelectron ionization energies are of the latter type. This is one reason why the X-ray emission line corresponding to an electron jumping from *nlj* to *n'l'j'* does not necessarily have *exactly* the same wavenumber as the difference between the two ionization energies, though a much more important problem in practice is that the X-ray absorption edge is influenced by the available empty orbitals accepting the electron coming from the inner shell.

The situation is far more complicated if the groundstate has positive S, or conceivably even when it is diamagnetic ($S = 0$) but not containing exclusively closed shells. The parameters of interelectronic repulsion separate the many levels of the ionized configuration containing two partly filled shells (Condon and Shortley, 1953). When the inner shell is an s orbital, the ionized system has the level having $(S + \frac{1}{2})$ below the other, $(S - \frac{1}{2})$, separated to the extent $(2S + 1)K(\mathrm{s}, \lambda)$ where $K(\mathrm{s}, \lambda)$ is the exchange integral of the two-electron operator between the s orbital and the λ orbital carrying the uncompensated spin density in the groundstate. To the first approximation (Jørgensen, 1962a, 1969a) this integral K is the average $\langle r^{-1} \rangle$ in atomic units of the product of two electronic densities $\psi_s^2 \psi_\lambda^2$. If one of the orbitals is highly exterior, such as the configuration 1s3d of the helium atom, the $K(1\mathrm{s},3\mathrm{d})$ is very small. On the other hand, in compounds containing a partly filled shell λ in the ground state, $K(\mathrm{s},\lambda)$ may be 0.5 to 1 eV. This was observed in the 1s signals of O_2 ($S = 1$) and NO ($S = \frac{1}{2}$) by the Uppsala group (Siegbahn *et al.*, 1969), the 3s signals of the $3d^3$ systems ($S = \frac{3}{2}$) chromium(III) (Helmer, 1973) and manganese(IV) (Fadley *et al.*, 1969), and of the $3d^5$ systems ($S = \frac{5}{2}$) manganese(II) and iron(III) (Fadley and Shirley, 1970a; Fadley *et al.*, 1969) and the 4s and 5s signals of the lanthanides (Cohen *et al.*, 1972). When an inner p shell is ionized, configurations such as $2p^5 3d^8$ containing 270 orthogonal states are fairly complicated and result in the observation of four 2p signals in high-spin ($S = 1$) nickel(II) complexes (Jørgensen, 1971a; Jørgensen and Berthou, 1972a), whereas low-spin ($S = 0$) nickel(II) complexes have only the two signals one would ascribe to $2p_{3/2}$ and $2p_{1/2}$ in closed-shell systems such as copper(I) and zinc(II). It is known from the X-ray spectra of NiO in absorption ($2p^6 3d^8 \rightarrow 2p^5 3d^9$) and of CuO in emission (Bonnelle and Jørgensen, 1964) that the intermediate coupling due to strong relativistic effects in the 2p shell distribute the energy levels formed by the sixty states of $2p^5 3d^9$ in two sets of each several eV wide, and the levels do not normally possess well-defined S. It has been argued by some authors that the structure in the 2p region of paramagnetic nickel(II) compounds has the same origin as the 2p satellites of copper(II) discussed below. This would necessitate a selection rule that such satellites lack significant intensity in diamagnetic compounds. However, an argument in favor of the interpretation as differing interelectronic repulsion is that the first and third 2p signal of ($S = 1$) nickel(II) is separated about 17.7 eV to be compared with the separation 17.2 eV in ($S = 0$) cases, and it is known (Jørgensen, 1972) that the chemical influence on such energy differences between two j levels is on the limit of the experimental uncertainty 0.1 eV. By the same token, the separation is close to 16.0 eV for cobalt(II) having $S = \frac{3}{2}$ but only 15.0 eV for cobalt(III) having $S = 0$ (Jørgensen and Berthou, 1972a; Frost *et al.*, 1972a).

The configuration with two partly filled shells $3d^9 4f^{q+1}$ has been theoretically studied by Sugar (1972a) and compared with X-ray absorption spectra of metallic lanthanum, erbium and thulium. Similar results (Sugar, 1972b) obtained for the configuration $4d^9 4f^{q+1}$ can be compared with the extended 4d regions in the photoelectron spectra of lanthanides (Jørgensen and Berthou, 1972a; Bonnelle *et al.*, 1972), *e.g.*, some 20 eV wide in ytterbium(III) to be compared with the distance 10.0 eV between the two Lu(III) signals. It is clear that the separations of energy levels due to differing interelectronic repulsions are quite large in configurations with two partly filled shells having small average radii. At the same time, the number of such energy levels is very high and frequently produces a broad hill of adjacent superposed photoelectron signals (Bonnelle *et al.*, 1972; Signorelli and Hayes, 1973).

We already mentioned the weak but numerous satellites in the photoelectron spectrum of gaseous neon (Siegbahn *et al.*, 1969) due to shake-up of the original ionized configuration $1s2s^2 2p^6$ to $1s2s^2 2p^5 3p$ or to shake-off to $1s2s2p^6$ loosing an additional electron. The corresponding weak satellites are difficult to detect in solids because of the much higher background though they have been discussed (Wertheim and Rosencwaig, 1971) in the case of rubidium salts. Stronger satellites occur in lanthanum(III) compounds and have been ascribed (Jørgensen and Berthou, 1972b) to interatomic shake-up corresponding to transfer of an electron from the orbital λ mainly localized on the ligands (having as a major component 2p in fluoride and oxides) to the empty orbitals 4f of the lanthanum ion. Hence, the ionized configuration producing the satellite contains three partly filled shells $\lambda^{-1} 3d^9 4f$ to be compared with $3d^9$ of the normal signals. The argument for this *electron-transfer satellite* character is that electron transfer spectra (Jørgensen, 1970) of cerium(IV) complexes show that the empty 4f shell is delocalized to a certain extent on the ligands. This is a necessary condition for a detectable intensity of the satellite and no longer occurs in heavy lanthanides. La(IV) lacking an inner electron such as 3d has a central field rather similar to groundstate Ce(IV) and comparable stabilization of the 4f orbitals. On the whole, the description of photoelectron spectra tends to be monoatomic and not consider interatomic processes in the condensed states of matter. The same is true for the spectra in the visible of ions containing a partly filled 4f shell (Jørgensen, 1973a) where the recognition of electron transfer spectra of Eu(III) and Yb(III) was retarded by the exclusive attention given to intra-atomic excitations. It is probable that the prominent satellites of the carbon 1s and oxygen 1s signals of carbon monoxide complexes of zerovalent chromium, iron, nickel, molybdenum, and tungsten (Pignataro, 1972) also are due to electron transfer in the ionized system from the d shell of the central atom, as known from "inverted electron transfer" bands of central atoms in low oxidation states in complexes of conjugated ligands which have low-lying empty MO's receiving an electron in the excited state (Jørgensen, 1970). The photoelectron signals of the central atoms do not show conspicuous satellites (Pignataro, 1972). However, an alternative explanation may be molecular shake-up corresponding to the strong absorption bands in the ultraviolet of such carbonyls. It is not possible to make this identification of the satellites of lanthanum(III) without taking into account the changed central field, since the optical electron transfer spectrum of La_2O_3 starts above 8 eV (Jørgensen, 1970), though conceptually, there is no difference between an electron-transfer satellite defined above and interatomic electron transfer shake-up energetically assisted by the charge separation effect. Madame Christiane Bonnelle (Jørgensen, 1973a) asks whether electron-transfer satellites might explain shoulders at *lower* I than the main 3d signals of Pr(III) and Nd(III). The regular evolution of electron transfer spectra suggests a decrease by 2.5 eV, so the satellites should at most coincide, but it is conceivable that the charge separation effects become more important.

The so-called *Theorem of Koopmans* is the statement (Koopmans, 1933) that the one-electron energies $-\varepsilon$ defined in appropriate way for Hartree–Fock wavefunctions are related by the simple relation $\varepsilon = I$ to the ionization energies I of the system if the other electrons do not modify their orbitals. The *rearrangement energy* is the difference between the Hartree–Fock energy for the *adapted* Ψ having slightly different orbitals for the remaining electrons and the original Hartree–Fock energy added to ε. As one would expect from the variation principle, the rearrangement energy is normally negative because the Koopmans assumption of *frozen orbitals* is an additional constraint. It is worthwhile here to mention the negative *correlation energy* defined (Löwdin and Yoshizumi, 1959) as the difference between the Hartree–Fock energy of a given state and its accurate nonrelativistic value allowing for mixing of electron configurations. In practice, the correlation energy is less negative for the ionized system than for the original system contributing to I being larger than ε whereas the rearrangement energy has the opposite, and usually predominant, effect. It is frequently stated that the correlation energy for N electrons is approximately $(1 - N)$ eV. However, the variation actually found is slightly steeper, and values known up to krypton suggest $-0.7Z^{1.2}$ eV for isolated atoms, proportional to the square root of the total energy close to $-13.6Z^{2.4}$ eV according to Gaspar (1967). In molecules, the amount of correlation energy is further increased by the sporadic difficulties for pure MO configurations which can vary strongly from one case to another. This is connected with the question of *relevant and irrelevant symmetry components* (Jørgensen, 1971b) and is a strong function of the internuclear distances. Thus, it is well known that the lowest MO configuration of the hydrogen molecule (Löwdin and Yoshizumi, 1959) grossly overestimates the energy for a large distance, suggesting the same probability for the ionic dissociation products H^- and H^+ as for the formation of two neutral atoms.

In this review, we concentrate on deviations from Koopmans' behavior of inner shells. The corresponding problem has been much discussed for delocalized MO's where it is suggested to multiply the calculated LCAO approximations to ε values by an empirical factor 0.92, that is a decrease of 8 percent (Brundle *et al.*, 1970). This correction represents a negative contribution from rearrangement and a positive contribution from correlation, as discussed above, but unfortunately it varies for the individual orbitals to such an extent that the order calculated for polyatomic molecules is not necessarily the same as the order of observed I values (Turner *et al.*, 1970; Siegbahn *et al.*, 1969). In this connection may be mentioned the method of Slater and Johnson (1972) named *Multiple-Scattering-Xα* which has been applied to molecules (Connolly *et al.*, 1973) such as CF_4, NH_3, and SF_6 for comparison with photoelectron spectra. This method gives better agreement than many highly elaborated approximations needing much longer calculations.

As far as inner shells go, the discussion can be divided into the monoatomic problem (as applicable to noble gases) and the question of interatomic relaxation effects. The monoatomic problem which has been the most carefully studied is the neon atom (Verhaegen *et al.*, 1971) where ε for the 1s orbital of the Hartree–Fock groundstate is 891.7 eV, 21.5 eV higher than the observed $I = 870.2$ eV. Of course, it can be argued that this decrease is only 2.4 percent, but still, it is almost as high as *I*(Ne2p). If an independent Hartree–Fock calculation is performed by Ne^+ with the constraint of having the electron configuration $1s2s^2 2p^6$, the energy is 868.6 eV higher than for the Hartree–Fock groundstate of the neutral atom. The agreement with the experimental value is almost perfect when the additional correlation effect between the two 1s electrons and their relativistic stabilization are taken into account. The naive conclusion from this coincidence is that the other nine electrons have the time to adapt their radial functions during the primary process of photoionization. Because the problems connected with the time scale discussed below do not allow a rapid rearrangement, the situation was rather paradoxical until Manne and Åberg (1970) resolved this apparent dilemma by pointing out that the many weak satellite signals due to shake-up and shake-off have an integrated intensity a quarter of the main signal. Since they occur between $I = 900$ and 1000 eV, the baricenter of the whole structure occurs at 886 eV, close to the ε value according to Koopmans. Said in other words, an electronic analogy to Franck–Condon's principle is in operation, but 80 percent of the structure is concentrated on the extreme side of low *I* corresponding to the configuration $1s2s^2 2p^6$ with the radial functions contracted relative to the neutral atom.

When the tables of Hartree–Fock ε values and photoelectron *I* values (to which a quantity close to 4 eV representing the Fermi level should be added) are compared (Siegbahn *et al.*, 1967) the deviations are not a definite percentage of *I* but are surprisingly near to the *square root* $0.8I^{0.5}$ eV corresponding to the combined rearrangement and differential correlation energy increasing from 8 eV for $I = 100$ eV to 25 eV for $I = 1000$ eV. On the whole, the deviations from Koopmans' behavior are slightly larger for the 2p orbitals in the 3d transition group than indicated by this empirical formula.

A numerically smaller, but far more unpredictable, relaxation effect is due to the surrounding atoms. Shirley (1972) demonstrated that the change of a noble gas atom by the absence of one electron in an inner shell is very similar to the presence of an additional proton on the nucleus. The calculated rearrangement energy for the 2p shell is 32 eV for krypton and 44 eV for xenon. *I* of $2p_{3/2}$ is 1678.4 and 4787.4 eV in the two cases. Unfortunately, it is not easy to obtain a reliable *I* for strongly cooled, solidified rare gases. *I* would be influenced in the opposite direction by the charging effect and by the interatomic relaxation.

Shirley (1972) argues convincingly that *I*(1s) of carbon, nitrogen, and oxygen in solid compounds is decreased almost 10 eV by interatomic relaxation. In particular, Mann has calculated the Hartree–Fock ε for the gaseous carbon atom to be 308.5 eV, and Shirley's perturbation treatment of the isolated atom would yield $I = 300$ eV. Since the carbon atoms in hydrocarbons are known to be almost neutral, and since I has been determined (Siegbahn *et al.*, 1969) to vary between 290.8 eV for methane and 290 eV for several heavier molecules, a chemical relaxation effect of 9–10 eV must be accepted. It had already been pointed out by Schwartz (1970a,b) that the discrepancy from Koopmans' theorem is 14 eV in methane and can be explained within 1 eV when a Hartree–Fock calculation is performed on CH_4^+ lacking the C1s electron. In a way, one may argue that this molecule has eight valence electrons, twice as many as the carbon atom, and that it is not surprising that the relaxation effect is twice as large. Solid graphite has I close to 288 eV after taking the Fermi level into account. It is not quite clear whether the decrease about 1 eV of I in benzene or other aromatic systems compared with aliphatic hydrocarbons is due to additional relaxation in the conjugated molecule or a slightly negative fractional charge on the carbon atoms.

A special case of irrelevant symmetry components is the question whether one should calculate ionization from a MO formed by linear combination of inner shells of different atoms of the same element in a given molecule. In a Hartree–Fock scheme, such I are lower than for a localized hole in one atom because a smaller residual positive charge is left on each atom. It is intuitively clear that the experimental I are not perceptibly changed by this effect (Murrell and Ralston, 1972) but one has to explain this subtle type of interatomic relaxation. One may argue from the time-dependent Schrödinger equation that the order of magnitude of migration time of a hole from one atom to another is $\hbar/E_{12}$ where E_{12} is the nondiagonal element of the Hamiltonian between the two localized states having identical energy. The fluorine 2s orbitals are unusual by having sufficiently strong interactions to show $I = 40.3$ eV and 43.8 eV for three degenerate and one MO in CF_4 and $I = 39.3$, 41.2, and 44.2 eV for two, three, and one MO in SF_6 (Siegbahn *et al.*, 1969; Connolly *et al.* 1973). However, truly inner shells have far weaker interactions, and the energy separations frequently given in literature are due to rounding-off errors. It can be argued (Jørgensen, 1972) the E_{12} has the order of magnitude of the overlap integral S_{12} times I, and if the exponential decrease of the radial function according to Schrödinger can be accepted, E_{12} between fluorine 1s orbitals in CF_4 and SF_6 is 10^{-8} eV corresponding to a migration time above 10^{-8} sec which is far from too long to produce any observable delocalization.

Jolly (1973) has proposed the compromise that the I values are determined by "half-ionized" cores in a state between the original groundstate and the fully ionized system after the photoelectron has left. It is true that intermediate states of relaxation may occur for sufficiently distant neighbor atoms for reasons to be discussed below, but the compromise does not seem plausible for the atom itself and the nearest neighbors. The point is that the signals observed in photoelectron spectra of solids are the main peaks in the sense of the 870.2 eV signal of neon, and they correspond to the fully adapted Ψ of the ionized system. If the signals were not such an extreme singularity in what is an electronic Franck–Condon structure according to Manne and Åberg (1970) they would be significantly broadened by representing a dynamic average of differing extent of relaxation. Since the total relaxation energy is 10–30 eV, the effective scattering of relaxation contributing to the width of the signal in solids is at most 1–5% in typical cases. The rational interpretation is 98% relaxation since the opposite extreme of very little relaxation is excluded by measurements of solid hydrocarbons.

Quite recently (Aarons *et al.*, 1973) careful estimates of relaxation energy between 12 and 22 eV for carbon, nitrogen, and oxygen 1s in gaseous molecules have been shown to scatter 2–3 eV for each element.

III. Hartree and Madelung Potentials, Electric and Chemical Polarizability of the Neighbor Atoms

However much the exchange integrals of the two-electron operator are important for practical calculations on atoms (Condon and Shortley, 1953) and molecules (Slater and Johnson, 1972) by decreasing the total amount of interelectronic repulsion in the antisymmetrized Slater determinant relative to a simple Hartree product, it is also true that a good picture including the major aspects of electrostatic interactions can be obtained by considering the *Hartree potential* U(x,y,z) (in monoatomic entities the *central field* U(r) with the nucleus at *origo*) consisting of the sum of nuclear attractions $-Z_1/r_1$ (we use the sign of the potential appropriate for electrons) and a positive contribution representing the repulsion by the *other* electrons. One advantage of this concept is that though the separation in nuclear and interelectronic quantities is feasible in small molecules, rather absurd divergences occur for large systems (Jørgensen, 1971b) where they are proportional to the fifth power of the linear extensions in the case of constant composition, whereas the number of electrons varies only as the third power. Another advantage is the treatment of *generated symmetries* of which the simplest case is the holohedrized symmetry U_{hol} and its residual counterpart, the hemihedrized U_{hem}:

$$\begin{aligned} U_{\text{hol}} &= \{U(x,y,z) + U(-x,-y,-z)\}/2, \\ U_{\text{hem}} &= \{U(x,y,z) - U(-x,-y,-z)\}/2. \end{aligned} \tag{2}$$

Another example involving formation of the average value U_0 on each spherical shell with constant r,

$$U = U_0(r) + U_{\text{oct}}(x,y,z) + U_{\text{res}}(x,y,z), \tag{3}$$

where an integration is needed, is followed by the octahedral symmetry U_{oct} generated (Jørgensen, 1971b) by the formation of the average value of $(U - U_0)$ in 48 points after the choice of the directions of the Cartesian axes. The point group O_h has the order 48, and quite generally, all the point groups can be generated by formation of average values.

In a polyatomic entity, the local central field $U_0(r)$ may be different from that prevailing in the neutral atom because of changed fractional charge of the atom. However, the surrounding atoms produce other effects. In molecular calculations, it is customary to talk about interatomic Coulomb effects, but a more general idea (Jørgensen, 1969a) is the *Madelung potential* $V(x,y,z)$ obtained as the sum over $-q_i/r_i$ where the atom with charge q_ie has the distance r_i from the point considered. It is important to remember the fact of electrostatics first pointed out by Newton in the analogous case of gravitation that the potential inside a uniformly charged sphere with the radius R and the charge q everywhere is q/R, and outside at a distance r from the center of the sphere, it is q/r as if all the charge were concentrated in a point. It cannot be argued that the Madelung potential is strictly constant inside a given atom in a compound, but another theorem states that spherically symmetric ions such as closed-shell systems or Mn(II) and Fe(III) in ^6S groundstates are perturbed exclusively by the spherically symmetric component $V_0(r)$. This is the reason why the Madelung energy of ionic crystals can be calculated as if the charges were points. The Madelung potentials in certain cubic crystals have the same absolute value α^*/a_0 but opposite sign of the cation and anion sites. If the unit cell parameter a_0 is given in Å, the constant α^* is 50.2 eV for the NaCl type, 29.3 eV for the CsCl type, 83.6 eV for the CaF_2 type, and 54.4 eV for the CuCl type, when the anions are univalent. If they carry a fractional charge q_i, these α^* values are multiplied by q_i, and in fully ionic bivalent cases, as might perhaps be represented by MgO,

ThO_2, and ZnS, α^* are twice as large. It is not generally true for all binary compounds (Jørgensen, 1969a) that the Madelung potential is the same with opposite sign for the cation and anion sites.

In crystalline alkaline halides, the Madelung potential varies from 12.5 eV in LiF to 6.4 eV in CsI. It was originally thought to be an argument for the I values without correction for charging effects (Jørgensen, 1971a) that they correspond closely to the values predicted from the ionization energies of the gaseous ions corrected with the Madelung potential. This is also true for MgO where $I(\mathrm{Mg2p}) = 57.7$ eV (the I' value corrected for charging effects is 53.1 eV) and $I(\mathrm{O2p}) = 14$ eV whereas the difference between the ionization energy 80.1 eV for gaseous Mg^{+2} and the Madelung potential 23.9 eV is 56.2 eV. Although an anion in free space cannot have a negative I, it is true that the I's determined for oxide from the Born–Haber cycle, or from a Hartree–Fock function constrained to contain six equivalent 2p electrons, are negative. Here, the estimate of O^{-2} would be -10 eV. On the other hand, this treatment is not perfect for beryllium(II) compounds where $I(\mathrm{Be1s}) = 153.85$ eV for gaseous Be^{+2}. We find $I = 123.1$ eV and $I' = 120.8$ eV for BeF_2, $I = 122.5$ eV and $I' = 119.6$ eV for K_2BeF_4, $I = 121.6$ eV and $I' = 119.3$ eV for BeO, and $I = 120.3$ eV and 118.5 eV for the carbonate complex $[Co(NH_3)_6]_2[OBe_4(CO_3)_6]$ (Jørgensen and Berthou, 1972a) and the order of magnitude of the Madelung potential is at most 30 eV in these compounds. For comparison, it may be mentioned (Siegbahn *et al.*, 1967) that $I^* = 112$ eV for metallic beryllium having a Fermi level close to 3 eV corresponding to $I \sim 115$ eV.

There is no doubt that the major difficulty for the Madelung description are cases where I for the anions are strongly increased from the values for the gaseous anions (F^- 3.40, Cl^- 3.61, Br^- 3.36, and I^- 3.06 eV) and approach or cross the values for the neutral atoms (F 17.42, Cl 12.97, Br 11.81, and I 10.45 eV). One may consider two kinds of asymptotic behavior with fractional anion charges either close to -1 or to 0. Actually, I for the loosest bound MO of gaseous halide molecules (Turner *et al.*, 1970) are remarkably close to the values (15.4, 12.1, 11.3, and 10.2 eV, respectively) obtained from the formula

$$I = (1 + 3.7\chi_{\mathrm{opt}})\mathrm{eV} \tag{4}$$

involving the *optical electronegativity* χ_{opt} derived from electron transfer spectra in the visible and the ultraviolet (Jørgensen, 1962a, 1970) where the reducing ligands have lost an electron in the excited state, and the central atom has decreased the oxidation number by one unit by accepting the electron in an empty or partly filled shell. For halide ligands, the constant 3.7 eV has been chosen in such a way that the Pauling values 3.9, 3.0, 2.8, and 2.5 are obtained, going from fluoride to iodide. In aqua ions, water has $\chi_{\text{opt}} = 3.5$ whereas it is 3.2 in sulfate and 2.6 to 2.8 in ligands bound by sulfur atoms to the central atom.

In many ways, it is rather surprising how closely similar are the I values predicted by a weak perturbation of the neutral atom and by a fully ionic situation including the Madelung potential (Jørgensen, 1969a; Jørgensen *et al.*, 1967). The importance of the Madelung potential for X-ray-induced photoelectron spectra was recognized as early as 1968 (Fadley *et al.*, 1968). A closer study of the alkaline halides (Jørgensen and Berthou, 1972a; Citrin and Thomas, 1972) established that the actual I' values corrected for charging effects are slightly lower than of corresponding gaseous halides described by Eq. (4) and usually 2 to 3 eV lower than calculated from the straightforward Madelung description. Thus, the predicted values would be $I(\text{Na2p}) = 38.37$ eV and $I(\text{Cl3p}) = 12.53$ eV for NaCl to be compared with observed $I' = 35.6$ and 10.0 eV (Jørgensen and Berthou, 1972a), or 36.4 and 10.9 eV (Citrin and Thomas, 1972).

It was already seen above that the total spreading of $I(\text{Be1s})$ in solid beryllium(II) compounds is 3 eV. Quite generally, almost all elements in the same oxidation state have intervals of I-values between 2 and 8 eV (Jørgensen and Berthou, 1972a). It is true that the solid fluorides presenting the highest I value tend to have the largest charging effects with the result that the interval of I' values usually is only two-thirds as large. There was a time (Fadley *et al.*, 1968) when it was believed that the main parameter determining the chemical shift is the oxidation state, and it is true that elements able to change their oxidation number by eight units have a particularly wide interval. Thus, $I(\text{N1s})$ in solid (Hendrickson *et al.*, 1969) and gaseous (Finn *et al.*, 1971) nitrogen compounds varies between 417 and 402 eV (409.9 eV for N_2), and $I(\text{S2p}_{3/2})$ varies from 180.4 eV for gaseous SF_6 (Siegbahn *et al.*, 1969) over 176 to 174 eV for various sulfates (Jørgensen and Berthou, 1972a) down to below 167 eV for certain sulfides and sulfur-containing complexes. Table I shows comparable results for perchlorates and chlorides. Nevertheless, it is not always possible to deduce the oxidation state from the chemical shift observed. Thus, Tl_2O_3 has *lower* I than all twenty thallium(I) compounds measured, PbO_2 lower I than all fourteen lead(II) compounds, and most cobalt(III) compounds have lower I than Co(II) (Jørgensen, 1971a; Jørgensen and Berthou, 1972a).

The obvious explanation of the I variation for a given oxidation state is a variation of the *fractional atomic charge* in transition group and post-transition group complexes. The high I values for fluorides and low I

values for sulfur-containing compounds are in agreement with the estimates derived from the visible spectra of central atoms containing from two to eight d electrons (Jørgensen, 1969a, 1971b) where the *nephelauxetic effect*, i.e., the decrease of parameters of interelectronic repulsion compared with the corresponding gaseous ion, can be brought in connection with the fractional atomic charge being below the oxidation number and with the character of the partly filled shell as nonbonding or slightly delocalized antibonding MO's. On the other hand, cyanides and iodides have higher I values than one would expect from the nephelauxetic effect.

If the central field $U(r)$ is modified by decreased fractional atomic charge (due to covalent bonding) and by the Madelung potential V_0 it is easy to see (Jørgensen, 1962a) that all the inner shells of the same atom ought to change I approximately to the same extent as verified by photoelectron spectra (Fadley *et al.*, 1968) and explaining the well-known fact of X-ray spectroscopy (Faessler, 1972) that the chemical shift of absorption edges is about ten times larger than of emission lines due to transitions between two inner shells. The argument of a definite change in $U(r)$ does not apply to the loosest bound valence electrons nor to a partly filled shell. This is clearly seen in the 4f group (Jørgensen and Berthou, 1972a; Jørgensen, 1973a) where the interval of chemical shifts for different compounds of M(III) is some 4 eV for I(M3d) and I(M4d) but only some 2 eV for I(M4f) whereas the increase of I from Ce(III) to Ce(IV), from Eu(II) to Eu(III), and from Tb(III) to Tb(IV) is about 12 eV for inner shells to be compared with the difference 9 eV of I(4f) between consecutive isoelectronic M(III) and M(IV). This dramatic change with one unit variation of the oxidation number is unique for the 4f group; it does not occur in uranium compounds.

The discussion above of relaxation effects indicates that a modified central field cannot be the whole truth. Actually, it cannot be excluded that intervals as broad as 5 eV are due to variations in the interatomic relaxation. Several authors have already pointed out (Fadley *et al.*, 1968; Citrin and Thomas, 1972) the possibility of electric polarization added to the Madelung potential expressed by Coulomb's law. The polarization might conceivably stabilize the groundstate of the system, increasing I, or it may favorize the ionization, decreasing I. Cubic crystals do not exhibit polarization due to electric dipoles. At sites belonging to the point group O_h the lowest multipole possessing total symmetry (Jørgensen, 1971b) is the 64 pole of very small practical importance. On the other hand, Citrin and Thomas (1972) discuss the polarization corresponding to electric dipoles induced around a hole left by the photoelectron. I is decreased about 2 eV in alkaline halides in agreement with previous estimates by solid-state physicists. Further on, a repulsion effect oscillating around 1 eV is introduced. The rationalization behind this quantity is the decreased ionic radius of the system having lost the photoelectron. With the usual Born expression, this decrease would release almost all the destabilization of the groundstate due to the core–core repulsion between adjacent ions. This repulsion compensates about one-tenth of the Madelung energy of the crystal (relative to gaseous ions) and may be connected with the kinetic energy operator (Jørgensen, 1971b). However, one may express doubts whether the sudden absence of one of the six equivalent p electrons in the loosest bound shell can be significantly represented by a decreased ionic radius, though such a development might enter the conventional Franck–Condon principle. It might also be argued that on an instantaneous picture, five-sixths of the core–core repulsion remains.

Actually, a quite different alternative is to suppose that neighbor atom electronic relaxation modifies the observed I values. As discussed in the next chapter, the more "polarizable" constituents in a chemical sense tend to induce low I values with the surprising exception of iodides and cyanides. A clear-cut example of influence of the environment has recently been reported (Citrin *et al.*, 1973) of noble gas atoms imbedded in metallic copper, silver, and gold. Thus, I(Xe3d) is decreased by 2.2, 2.0, and 2.0 eV in the three metals, krypton and argon show slightly larger effects, and I(Ne1s) of neon in copper is decreased 3.0 eV. Many authors have tried to establish linear relations between I and the fractional atomic charge. Such attempts meet two major difficulties: that the linear relation based on a definite model of LCAO type may have a grossly incorrect slope (*e.g.*, if all the charges are underestimated by neglecting the influence of the Madelung potential) and that relaxation effects may imitate fractional atomic charges (which are not expected to be prominent on neon atoms in a metal). Noble gases in coinage metals are further discussed by Citrin and Hamann (1973).

IV. Experimental Data and the Softness Scale According to Pearson

For the chemist, it is rather appalling that the typical range of ionization energies of an inner shell varying in different compounds of the same element rarely is below 3 eV and frequently above 10 eV when it is realized that the strongest single bonds have dissociation energies close to 5 eV or 115 kcal/mole. By the same token as the X-ray spectra are said in school to be a characteristic property of a definite element, it is felt that the inner shells have no reason to be so sensitive to chemical behavior. However, it must be stressed that if the changes of I only are indicators of changes in the prevailing Hartree potential, they do not enter to the first approximation in the total energy because the effects of interelectronic repulsion are counted twice (Jørgensen, 1962a) in the sum over the one-shot ionization energies considered here, which should not be confused with the larger sum of consecutive ionization energies. The situation is different with genuine chemical effects on deep-lying orbitals as known from fluorine 2s in CF_4 and SF_6 (Jørgensen, 1972) producing an overall destabilization like the core–core repulsion in crystalline NaF.

Table I gives $I(\mathrm{Cl}2\mathrm{p}_{3/2})$ and $I(\mathrm{Cl}3\mathrm{p})$ for some perchlorates and chlorides selected among 90 chlorine compounds (Jørgensen and Berthou, 1972a). When the correction for charging effects mentioned in the introduction can be performed, the corresponding I' values are also given. Table II gives $I(\mathrm{K}2\mathrm{p}_{3/2})$, $I(\mathrm{K}3\mathrm{s})$, and $I(\mathrm{K}3\mathrm{p})$ for a series of potassium salts. Table III gives $I(\mathrm{M}4\mathrm{d})$ and $I(\mathrm{M}4\mathrm{f})$ for ytterbium(III), lutetium(III), hafnium(IV), and tantalum(V) compounds. Finally, Table IV gives $I(\mathrm{Th}4\mathrm{d})$ and $I(\mathrm{Th}4\mathrm{f})$ for thorium(IV) compounds.

On the average, the perchlorates have $I(\mathrm{Cl}2\mathrm{p}_{3/2})$ 9 eV higher than the

TABLE I

PHOTOELECTRON SIGNALS OF PERCHLORATES AND CHLORIDES (eV)

	C'_{st}	C_{st}	$I(\mathrm{Cl}2\mathrm{p}_{3/2})$	$I'(\mathrm{Cl}2\mathrm{p}_{3/2})$	$I(\mathrm{Cl}3\mathrm{p})$	$I'(\mathrm{Cl}3\mathrm{p})$
$KClO_4$	—	4.8	215.0	—	—	—
$[Co(H_2O)_6](ClO_4)_2$	—	4.8	214.5	—	—	—
$[PtNH(C_2H_4NH_2)_2Cl_3]ClO_4$	—	4.5	213.7,204.5	—	—	—
$[Ni(NH_2C_2H_4NH_2)_2](ClO_4)_2$	—	4.6	213.3	—	—	—
$[N(C_4H_9)_4]ClO_4$	—	4.8	212.6	—	—	—
CsCl	0.7	5.2	207.3	202.8	13.8	9.3
Cs_2PtCl_6	1.6	4.3	206.9	204.2	15.7	13.0
Rb_2PtCl_6	2.5	5.0	206.8	204.3	15.6	13.1
NaCl	3.0	5.2	206.1	203.9	12.2	10.0
RbCl	2.8	5.2	205.7	203.3	11.9	9.5
KCl	2.3	4.8	205.5	203.0	11.9	9.4
Na_3RhCl_6, $12H_2O$	—	5.0	204.4	—	12.4	—
K_2ReCl_6	—	3.7	204.2	—	12.1	—
$[N(C_4H_9)_4]_2PtCl_4$	—	5.3	203.5	—	11	—
$[Ru(NH_3)_6]Cl_3$	—	4.7	203.0	—	—	—
$[As(C_6H_5)_4]Cl$	—	6.2	202.0	—	—	—

corresponding chlorides, a clear effect of increasing the oxidation state from Cl(−I) to Cl(VII). In agreement with the Madelung description, the larger cations induce lower I in the perchlorates. The same is to a certain extent true for chlorides, but the I' values corrected for charging effects show a somewhat different order. The rather covalent $PtCl_6^{-2}$ has a less negative charge on the chlorine atoms. It is particularly clear that the I' values between 10.0 and 9.3 eV influenced by the Madelung potential in the alkaline halides have increased to 13 eV in $PtCl_6^{-2}$ higher than 12 eV found for most gaseous (and rather covalent) halides (Turner *et al.*, 1970) and 12.1 eV derived from Eq. (4). On the other hand, the high I value for CsCl seems to be due to charging effects; I' is quite moderate. It is not possible to find MO's corresponding mainly to Cl3p in perchlorates where five different I values in the 11 to 20 eV region represent orbitals predominantly constituted by oxygen 2p in the LCAO approximation. For group-theoretical reasons, one would expect seven I values (Jørgensen, 1970) from MO's constituted mainly by chlorine 3p in $PtCl_6^{-2}$, and the entry in Table I indicates an average position weighted by the relative signal intensities. In gaseous SF_6 these I values are fairly clearly resolved (Siegbahn *et al.*, 1969) whereas the isoelectronic PF_6^- and SiF_6^{-2} in solids still show some residual structure in the fluorine 2p region (Jørgensen *et al.*, 1972). The differences in excitation energy derived from electron transfer spectra of hexahalide complexes (Jørgensen, 1970) such as $IrCl_6^{-2}$ and $PtCl_6^{-2}$ produce an interval 2 eV wide.

It is striking (Jørgensen and Berthou, 1972a; Jørgensen *et al.*, 1972) that the variation of $I(K2p_{3/2})$ (cf. Table II) follows very closely the order of Pearson (1963, 1966, 1968) of the anions as *hard* and *soft* bases, with the exception of iodide (though the I' value is more in the middle of the series). Many chemists criticize these new words originally intended to convey the thought that hard central atoms and other Lewis acids (or antibases, as J. Bjerrum proposes) such as Al(III), La(III), and Th(IV) preferentially bind hard bases such as fluoride and oxygen-containing ligands, whereas soft central atoms such as Cu(I), Pd(II), Ag(I), Au(I), Au(III), Hg(II), and Tl(III) prefer bromide, iodide, phosphines, and sulfur-containing ligands. There is no doubt that this concept is not entirely new and is based, in part, on the geochemical classification of Goldschmidt in lithophilic and chalkophilic elements and on the old analytical techniques using H_2S for separations. Nevertheless, it seems to have certain quantitative connotations. Ahrland (1968) proposed the softness parameter

$$\sigma_A = (1/z)[(I_1 + I_2 + \cdots + I_z) + H_z] \tag{5}$$

for the cation with charge $+z$ derived from the sum of consecutive ionization energies of the gaseous ion and the (negative) heat of hydration H_z of the M^{+z} when transferred from *vacuo* to dilute aqueous solution. The parameter Eq. (5) varies from 0.0 eV for Li(I), 1.6 eV for Al(III), 2.28 eV for H(I), 4.1 eV for Ag(I) to 4.6 eV for Hg(II). If the theory of *hydration energy differences* (Jørgensen, 1969a, b) can be applied, where the chemical ionization energy for aqua ions increasing the oxidation number from $(z - 1)$ to $+z$ is the sum of 4.5 eV and the standard oxidation potential relative to the hydrogen electrode (entropy contributions can be neglected for this purpose), then the difference between I_z for the corresponding gaseous ion and the chemical ionization energy can be written $(2z - 1)\kappa$ where κ is a parameter turning out to be about 5.6 eV for Be(II) and Al(III), 5.3 eV in the 3d group, 4.3 eV in the 4f group, and 3.9 eV in the 5f group. This is another way of saying that the heat of hydration H_z is $-z^2\kappa$. In this approximation, Eq. (5) reduces to

$$\sigma_A = \frac{1}{z}(I_1 + I_2 + \cdots + I_z) - z\kappa. \tag{6}$$

It is easy to generalize Eq. (5) to anions, and for the halides $\sigma_A = -8.7$ eV for F^-, -7.5 eV for Cl^-, -6.9 eV for Br^-, and -6.1 eV for I^-.

It cannot be argued that the variation of the potassium ionization energies in Table II is due to a fractional atomic charge much below $+1$. The opposite argument based on the concept of equilibrated electronegativities neglects the importance of the Madelung potential (Jørgensen, 1962a). If the observed variation was due to the Madelung potential alone, I of the cation would be the lowest when the anions are the smallest. However, both in the series of I and of I' the monoatomic halides are in the middle, surrounded by higher values for fluorine-containing anions, and oxygen-containing anions in the middle, and with large complex anions of noble metals with heavier halides and pseudo-halides (CN^- and both N- and S-bound thiocyanate) and selenocyanate itself at the lowest positions. The position of the tetraphenylborate is quite interesting; it becomes conducting under the influence of the X rays and must be rather high up on the list of I' values only slightly below 298.9 eV. Nevertheless, if this value is 298.4 eV, it is lower than I' of the four salts of discrete, monomeric fluorine-containing anions (whereas $KNiF_3$ is a cubic perovskite with each fluoride bound to two colinear nickel atoms). By far the most plausible diagnosis is that the interatomic relaxation effects are less pronounced in Pearson's hard anions (such as the fluorine-containing complexes) than in the case of soft anions, whether it be large polyatomic

TABLE II

PHOTOELECTRON SIGNALS OF POTASSIUM SALTS (eV)

	C'_{st}	C_{st}	$I(K2p_{3/2})$	$I'(K2p_{3/2})$	$I(K3s)$	$I(K3p)$
K_2BeF_4	2.0	5.0	301.5	298.5	41.6	25.4
K_2SiF_6	2.4	5.0	301.4	298.8	41.9	25.5
KPF_6	2.7	4.5	301.0	299.2	39.0	24.8
KBF_4	2.8	5.4	301.8	299.2	41.4	25.7
KI	1.6	4.7	301.0	297.9	40.7	24.8
$KNiF_3$	1.3	4.5	300.9	297.7	(38)	25.2
KIO_4	2.8	5.3	300.5	298.0	40.1	24.3
KF	1.7	4.1	300.3	297.9	40.3	24.1
KCl	2.3	4.8	300.3	297.8	40.2	23.9
KIO_3	3.2	5.3	300.2	298.1	40.5	23.8
$KOsO_3N$	2.1	4.4	300.0	297.7	40	24.0
KBr	3.1	4.6	299.8	298.3	39.3	23.3
$KAg(CN)_2$	2.7	4.5	299.3	297.5	—	—
$KB(C_6H_5)_4$	—	5.1	298.9	—	39.0	22.7
K_2ReCl_6	—	3.7	298.7	—	39.0	22.4
$K_4Fe(CN)_6$	—	5.1	298.6	—	39.0	22.8
$K_3Cr(NCS)_6$	—	4.8	298.4	—	38.7	22.6
$KMnO_4$	—	5.1	298.4	—	38.5	22.4
$K_2Pt(SCN)_6$	—	4.8	298.3	—	38.5	22.5
KSeCN	3.3	5.0	298.2	296.5	39	—
$K_3Fe(CN)_6$	—	5.4	297.9	—	38.1	21.9

entities or $SeCN^-$. Hence, the photoelectron spectra have provided a measurable expression for an effect strongly appealing to chemical intuition. It must be emphasized that though the molar polarizabilities derived from refractive indices are approximately additive (Salzmann and Jørgensen, 1968; Jørgensen, 1969c) they are highly different from the values for gaseous ions, and they have only a very remote relation to chemical softness.

It is necessary to discuss the case of potassium iodide in more detail. The Madelung potential at the potassium site is +9.4 eV in KF and +7.1 eV in KI. Hence, one would expect I in the former compound to be 2.3 eV lower. The observed decrease in I is 0.7 eV whereas the observed I' values are identical. Citrin and Thomas (1972) find differences in polarization energy 0.83 eV for NaF and NaI and 0.66 eV for CsF and CsI, suggesting about one-third of the difference in Madelung energy is compensated. The fact that all of this difference is cancelled in the I' values suggests an additional effect. Though the static polarization of the groundstate of NaCl-type crystals is of minor practical importance, it may still be true that a kind of van der Waals stabilization of the groundstate is destroyed by the ionization. It is clear that a quantitative version of such an idea soon introduces more parameters than known data. Nevertheless, it is quite characteristic that the $I'(\mathrm{K2p}_{3/2})$ of the salts of BF_4^- and PF_6^- both are 1.3 eV higher than of F^-. Since the effective ionic radius of the two former complexes is about twice that of fluoride, it would not be surprising if the Madelung potential is 3 eV less positive in the two former salts. Again, the expected increase of I is partly cancelled by an unknown effect. It is conceivable that the relaxation in the hexafluorophosphate is more pronounced, and even more so in $KNiF_3$ where each K^+ is surrounded by twelve F^- and not six F^- as in KF.

In Table III, the structure of four to six broad signals in the Yb4d region corresponds to ionization to several of the 140 states of the configuration $4d^9 4f^{13}$ (writing only the partly filled shells) which must be distributed over more than 21 eV in a situation of intermediate coupling, since the separation 10.0 eV between the two signals of lutetium and 10.7 eV of hafnium simply represents the relativistic (spin-orbit) splitting between $4d_{3/2}$ and $4d_{5/2}$ at lower I. The configuration $4d^9 4f^{13}$ is also the excited states of strong absorption bands in the soft X-ray region of $4f^{12}$ thulium(III), but here the intensity distribution is determined by selection rules of electric dipole radiation (Sugar, 1972b). Wertheim *et al.* (1971) were the first to point out that whereas the $4f^7$ system GdF_3 has one signal due to I(Gd4f) slightly above I(F2p) the subsequent $4f^8$ to $4f^{13}$ possess two signals due to I(M4f). This can readily be explained by the possibility in the second half of the 4f group of the groundstate changing the total spin-quantum number S to $(S + \frac{1}{2})$ having less interelectronic repulsion than states with $(S - \frac{1}{2})$. This argument can be extended to the general theory of spin-pairing energy in the 4f group (Jørgensen, 1969a, 1973a) and actually, the separation 3.7 eV between the two Yb4f signals is larger than the separation $2D \sim 1.7$ eV between the baricenters of the triplet and singlet states of $4f^{12}$ but represent a weighted average of two sets of excited levels. The distribution of photoelectron signal intensity can be described by coefficients of fractional parentage, and Dr. P. A. Cox has calculated such coefficients of all numbers of f electrons in good agreement with our spectra (Cox *et al.*, 1973). Thus, (54/91) of the intensity concentrated at lower I (showing doublet structure under favorable conditions of high resolution with a high-intensity magnesium anticathode) corresponds to the terms ^3H and ^3F of $4f^{12}$ whereas (22/91) of the intensity is concentrated

TABLE III

PHOTOELECTRON SIGNALS OF YTTERBIUM, LUTETIUM, HAFNIUM, AND TANTALUM COMPOUNDS (eV)

	C'_{st}	C_{st}	I(M4d)	I'(M4d)	I(M4f)	I'(M4f)
YbF_3	2.2	4.7	216,209,202,194.7	213,206.5,199.5,192.2	20.5,16.7	18.0,14.2
$Yb(IO_3)_3$	2.0	4.7	214,207.3,199.2,193.9	211.204.6,196.5,191.2	20,16.0	17,13.3
$YbTaO_4$	1.9	4.7	207,200,193.3	204,197,190.5	19.4,15.7	16.6,12.9
$Yb_2Hf_2O_7$	2.2	4.7	208,200,193.1	205,197.5,190.6	18.7,15.0	16.2,12.5
Yb_2O_3	3.9	5.4	213,205.9,199.6,191.9	212,204.4, 198.1,190.4	17.9,14.2	16.4,12.7
$YbVO_4$	2.8	4.9	205,199,192.2	203,197,190.1	18.2.14.3	16.1.12.2
LuF_3	0.3	5.5	218.6,208.7	213.4,203.5	(22),20.5	(17),15.3
LuF_3, repeated	0.4	4.9	217.6,207.6	213.1, 203.1	(21),19.7	(17),15.2
$Lu(IO_3)_3$	2.0	4.5	214.8,204.8	212.3,202.3	16.8	14.3
Lu_2O_3	2.7	4.8	213.9,203.9	211.8,201.8	16.0	13.9
K_2HfF_6	2.7	5.0	232.9,222.4	230.6,220.1	(27.5),26.0	(25.2),23.7
$Hf(IO_3)_4$	2.6	5.0	232.6,221.9	230.2,219.5	(27.3),25.6	(24.9),23.2
HfO_2	2.2	4.5	231.5,220.8	229.2,218.5	(26.5),25.0	(24.2),22.7
$Yb_2Hf_2O_7$	2.2	4.7	231.4,220.7	228.9,218.2	(26.1),24.4	(23.6),21.9
$Hf(C_6H_5CHOHCO_2)_4$	—	5.0	230.2,219.5	—	(24.9),23.2	—
K_2TaF_7	3.1	4.5	250.7,239.2	249.3,237.8	37.5,35.7	36.1,34.3
$YbTaO_4$	1.9	4.7	250.1,238.3	247.3,235.5	36.3,34.6	33.5,31.8
Ta_2O_5	2.8	4.9	248.8,237.4	246.7,235.3	35.1,33.3	33.0,31.2

on the coinciding ^3P and ^1I. It is known from the absorption spectra of thulium(III) compounds (Jørgensen, 1962a) that these two terms are situated 4.5 eV above the groundstate $^3H_6$ which itself is 1.0 eV below the baricenter of the six levels belonging to ^3H and ^3F. The 4f signals of Lu(III) are slightly asymmetric and under conditions of high resolution show a shoulder at higher I. The origin of this splitting which is somewhat more pronounced in Hf(IV) and Ta(V) is the separation between $4f_{5/2}$ and $4f_{7/2}$ known to be 1.3 eV in all Yb(III) compounds (Jørgensen, 1962a). Dr. Jack Sugar was so kind as to indicate this separation to be 1.46 eV in gaseous Lu^{+4}. It increases smoothly to 5.3 eV in bismuth ($Z = 83$) as a function of increasing atomic number (Jørgensen and Berthou, 1972a; Jørgensen, 1972).

The thorium(IV) compounds included in Table IV exemplify the typical chemical effects on I values of inner shells of an element with the same oxidation state. It is noted how the separation 9.35 eV between $4f_{5/2}$ and $4f_{7/2}$ remains invariant (Jørgensen, 1972). The apparent scattering around 37.5 eV of the separation between $4d_{3/2}$ and $4d_{5/2}$ is probably due to experimental uncertainty; the background is intense and the signals broad. Like for hafnium(IV), I' cannot be determined for the mandelate. Among the organic ligands, the acetylacetonate also has low I values whereas the oxalate is surprisingly high. In Table IV, I' has only been given for the more intense signals $4d_{5/2}$ and $4f_{7/2}$.

The four tables are only intended to exhibit a few typical examples among the more than 600 solid compounds recently studied (Jørgensen and Berthou, 1972a). They show that though a description based on fractional atomic charges and Madelung potentials discloses a major part of the truth, undoubted effects above 2 eV are due to relaxation in neighbor atoms following chemical ideas of polarizable atoms, which is not exactly the same concept as the electric dipolar polarizability. For Pearson (1963) metals are intrinsically "soft" and a corollary of the idea of interatomic relaxation is that gaseous atoms should have higher I values than liquid and solid metals. If europium atoms (Fadley and Shirley, 1970a) really have $I(4f) = 12.0$ eV, this is much larger than $I = 4.6$ eV for metallic europium. These values may be compared with $I = 8.9$ eV and $I' = 6.9$ eV for $EuSO_4$ and $I = 14.9$ eV and $I' = 12.0$ eV for Eu_2O_3. Since the standard oxidation potential of europium(II) aqua ions relative to the hydrogen electrode is -0.4 V, the chemical ionization energy I_{chem} is only 4.1 eV. However, it is important to remember that chemical equilibrium refer to adiabatic reactions where the internuclear distances are allowed to rearrange.

TABLE IV

PHOTOELECTRON SIGNALS OF THORIUM COMPOUNDS (eV)

	C'_{st}	C_{st}	$I(Th4d_{3/2})$	$I(Th4d_{5/2})$	$I'(Th4d_{5/2})$	$I(Th4f_{5/2})$	$I(Th4f_{7/2})$	$I'(Th4f_{7/2})$
ThF_4	1.3	4.9	723.5	686.1	682.5	354.75	345.4	341.8
$Th(C_2O_4)_2$	1.9	4.6	722.6	685.1	682.4	353.45	344.1	341.4
$Th(IO_3)_4$	2.2	4.7	722.4	684.3	681.8	352.85	343.5	341.0
ThO_2	1.0	4.6	720.6	683.4	679.8	352.0	342.65	339.05
$[Co(NH_3)_6]_2[Th(CO_3)_5]$	2.3	4.2	720.7	683.0	681.1	351.8	342.45	340.55
$Th(CH_3COCHCOCH_3)_4$	2.3	4.3	720.0	682.5	680.5	351.1	341.75	339.75
$Th(C_6H_5CHOHCO_2)_4$	—	4.3	719.7	682.4	—	350.85	341.5	—

The distances between different $4f^7$ terms formed by ionization of metallic terbium are 12% larger than in the ultraviolet spectrum of the isoelectronic Gd(III). This increase is about 6 percent in ionized thulium compared with the $4f^{11}$ levels of erbium(III) (Cox *et al.*, 1973). The increase is not as large as 20% expected for gaseous M^{+4} and may indicate a nephelauxetic effect (Jørgensen, 1969a) in the adapted radial function of the ionized system.

V. What Are One-Electron Energies in Ligand Field Theory?

Bethe suggested in 1929 the electrostatic model of the ligand field theory (Griffith, 1961; Wybourne, 1965): that the low-lying energy levels of d and f group complexes containing a partly filled shell are determined by the nonspherical component $V - V_0(r)$ of the Madelung potential divided in various parts like the Hartree potential in Eq. (3). Since this model was remarked by the chemists after 1951, it soon turned out to be numerically inadequate, and it was realized that the main effect is the antibonding character of some of the five d or seven f orbitals. The *angular overlap model* was originally (Jørgensen *et al.*, 1963) intended to describe the very small splittings (Wybourne, 1965) of each J-level of $4f^q$ in compounds as moderate energetic separations (below 0.1 eV) of the seven 4f orbitals, but it was extended (Schäffer and Jørgensen, 1965) to d-group complexes where the typical energy separations between the subshells each consisting of one, two, or three of the five d orbitals are 1 to 5 eV. It is possible (Schäffer, 1973) to establish mathematical relations between the antibonding character proportional to the square of the overlap integrals S_{MX} between central atom and ligand orbitals and the parameters of the electrostatic model of the ligand field if one refrains from calculating them from the series expansion of V.

The development of ligand field theory since 1955 has brought the subject (Jørgensen, 1971b) in close contact with the semiempirical MO treatment of heteronuclear molecules though a characteristic emphasis on interelectronic repulsion in partly filled shells has remained. The Hückel model of homonuclear molecules reduces to a topological matrix if only one kind of nearest neighbors is considered and represented by a nondiagonal element β. In such a case, it is convenient to write the diagonal element α common to all the orbitals as a factor times β. In mildly heteropolar cases, such as carbon and nitrogen atoms in pyridine, it is reasonable to introduce diagonal elements α_C and α_N as two different constants multiplying β. The influence of accumulating fractional atomic charges on

α of a given atom was introduced by Streitwieser as the ω method. Nevertheless, the description of typically inorganic compounds with large separations is only possible when considering also the Madelung potential (Jørgensen *et al.*, 1967).

Until 1971, it was believed that the physical significance of the angular overlap model is that nonbonding orbitals of the partly filled shell of the central atom have the diagonal energy H_M (containing contributions independent of the symmetry type of the orbital such as the full Madelung potential) whereas the antibonding orbitals have their energy increased by the second-order perturbation $(H_{MX} S_{MX})^2/(H_M - H_X)$ assuming the diagonal element H_X of the ligand orbitals to be considerably more negative than H_M and assuming the Hückel nondiagonal β to be proportional to the overlap integral S_{MX}. This general idea can be expressed in many specific models where Wolfsberg and Helmholz suggested H_{MX} to be the arithmetic average $(H_M + H_X)/2$ multiplied by a constant somewhere between 1.5 and 2. One advantage of this expression is that the diagonal sum rule would be valid (Jørgensen, 1962a) if the constant was 1, whereas a larger constant ensures the fact known from molecular spectroscopy (Herzberg, 1966) that the antibonding MO's are twice or thrice as antibonding as the corresponding bonding MO's are bonding. This fact is also the origin of the core–core repulsion preventing the implosion of molecules and crystals. Other authors such as Basch, Viste, and Gray (1966) use H_{MX} equal to the geometric average $-(H_M H_X)^{1/2}$ with the practical appeal for computers that the effect of a given orbital can be annihilated by putting its diagonal element equal to zero. A problem turned up early (Fenske *et al.*, 1966) that anion complexes such as CrF_6^{-3} have positive eigenvalues (corresponding to spontaneous loss of electrons *in vacuo*) unless a background potential is introduced, representing the Madelung potential from surrounding cations, or a dielectric hydration or other solvation energy.

However, one of the fundamentals of such models has been removed by the photoelectron spectra of solid 4f group compounds (Jørgensen, 1973a; Wertheim *et al.*, 1971) showing *higher* I(M4f) in most cases than the loosest bound orbitals of the ligands representing, for instance, I(F2p) or I(O2p). It is seen in Table III that the lowest ionization energy 4f of ytterbium(III) by a chemical shift follows about 2 eV *above* the loosest bound ligand orbitals. It cannot be argued that the signals have been incorrectly assigned since they are highly different in Yb(III) and Lu(III) and absent in Y(III) and also because of the general proportionality (Jørgensen and Berthou, 1973) between signal intensities and the number of electrons times their average value of r^{-2}. The same situation occurs in some, but not all, 3d group complexes. Gaseous VCl_4 has a weak signal (Cox *et al.*, 1970) at 9.4 eV corresponding to the single 3d-like electron, whereas a set of signals starting at 11.8 eV are closely similar to those of $TiCl_4$ and correspond to MO's mainly formed from Cl2p. However, the situation is different in volatile (Evans *et al.*, 1972, 1973) tris(hexafluoroacetylacetonates) $M(CF_3COCHCOCF_3)_3$ having a standard set of MO signals between 10 and 18 eV in the closed-shell systems M = Al, Sc, and Ga whereas, in agreement with general expectations from ligand field theory, a separate signal occurs at $I = 7.9$ eV for M = Ti containing one, 8.7 eV for V with two, and $I = 9.6$ eV for Cr with three d-like electrons. It also agrees with the increasing oxidizing character across the transition group and the decreasing excitation energy of electron transfer spectra (Jørgensen, 1970) that such a change of 0.8 eV is observed between consecutive elements. However, for M = Fe and Co, the signals belonging to the d-like orbitals are hidden around 11 eV in the structure due to the delocalized MO. It must be noted that the conditions of intensity are almost opposite (Evans *et al.*, 1973) for photoelectron spectra induced by 21.2 and by 1486.6 eV photons. In the former case, the d-orbitals give weak signals, and it is almost impossible to detect 4f signals though (Evans *et al.*, 1973) cerium(III) methylcyclopentadienide $Ce(C_5H_4CH_3)_3$ seems to show one at 7.3 eV.

The hexafluoroacetylacetonates contain the chromophore $M(III)O_6$ like the aqua ions and ruby (Jørgensen, 1963a) with approximately octahedral symmetry. Hence, two of the five d-orbitals are strongly antibonding and form a subshell at some 2 eV higher energy (as far goes the optical excitations discussed in ligand field theory) than the lower subshell consisting of three orbitals. It is quite clear that the condition for a well-defined oxidation state (Jørgensen, 1969a) determined from the number of d-like electrons is not that I_M is lower than I_X but that the electron affinity of the partly filled shell is smaller than any of the I_X. Hence, there is nothing inconsistent in shells with small average radii and consequent huge parameters of interelectronic repulsion having much lower electron affinity than ionization energy. However, the case of the d group complexes discussed here can be stated in a much more provocative way for the high-spin ($S = 5/2$) iron(III) compound having an electron in each of the five d-like orbitals (whereas the cobalt(III) compound has all the six d-like electrons in the lower subshell) that sometimes, an *antibonding MO has a higher ionization energy than the bonding counterpart of the same symmetry type.*

This apparent paradox can be resolved by a closer study of the effects of *interelectronic repulsion.* It has frequently been stated (Jørgensen, 1968, 1969a, 1971b) that interelectronic repulsion is the key to the understanding of the behavior of the transition groups, but emphasis has usually been laid on the increasing energy for decreasing S corresponding to one of Hund's rules. The expression for spin-pairing energy containing the contribution $-DS(S+1)$ (where D is about 0.8 eV in the 3d and 4f groups and about 0.4 eV in the 4d, 5d and 5f groups) explains many features of the 4f group (Jørgensen, 1973a) and the application of Eq. (4) to predict I(M4f) of M(III) compounds from what is technically (Nugent *et al.*, 1971) the uncorrected optical electronegativity of M(IV). It also explains why it is frequently an advantage for a 3d complex to be high-spin with a large S value and less interelectronic repulsion even when the condition is the presence of antibonding electrons as in most Mn(II) and Fe(III) compounds.

However, the J-integrals expressing the difference (Jørgensen, 1962a, 1969a) between ionization energy and electron affinity are about ten times larger than the spin-pairing parameter D. For our purposes, it is a sufficient approximation to write the energy of the electron configuration $a^m b^n$

$$-m\mathrm{W}(a) - n\mathrm{W}(b) + \frac{m(m-1)}{2}\,\mathrm{J}(a, a) + mn\mathrm{J}(a, b) + \frac{n(n-1)}{2}\,\mathrm{J}(b, b), \qquad (7)$$

where W are the ionization energies of a system containing only one electron, and the integral J(a, b) representing the Coulombic interaction between the charge densities $\psi_a{}^2$ and $\psi_b{}^2$ is about $\langle r^{-1}\rangle$ hartree for the most extended orbital a or b if they are highly different. In the opposite case of two electrons in the same shell having $\langle r^{-1}\rangle$ measured in reciprocal bohr units, the J(a, a) is approximately $0.65\langle r^{-1}\rangle$ hartree. Equation (7) is readily extended to more than two electrons in the same shell. It is also possible (Jørgensen, 1969a) to consider W and J of Eq. (7) as phenomenological parameters of interelectronic repulsion to be deduced from observed configuration baricenters, rather than calculated for Hartree–Fock functions. Gaseous atoms and monoatomic ions formed by transition group elements are quite successfully described (Jørgensen, 1973b) by such parameters.

As a specific example of Eq. (7) for our purposes, Dr. John Ammeter has proposed to consider a system with at most three electrons choosing between a d and an s shell. The configurations of interest to us are

$$\begin{aligned} \mathrm{E}(\mathrm{ds}^2) &= -\mathrm{W(d)} - 2\mathrm{W(s)} + 2\mathrm{J(d, s)} + \mathrm{J(s, s)}, \\ \mathrm{E}(\mathrm{d}^2\mathrm{s}) &= -2\mathrm{W(d)} - \mathrm{W(s)} + \mathrm{J(d, d)} + 2\mathrm{J(d, s)}, \\ \mathrm{E(ds)} &= -\mathrm{W(d)} - \mathrm{W(s)} + \mathrm{J(d, s)}, \\ \mathrm{E}(\mathrm{s}^2) &= -2\mathrm{W(s)} + \mathrm{J(s, s)}. \end{aligned} \qquad (8)$$

It can be seen from the neutral scandium atom that the baricenter of the configuration $3d4s^2$ almost coincides with the groundstate, whereas the baricenter of the 90 states belonging to $3d^24s$ is situated at 2.04 eV above the groundstate. The ionization energy of Sc forming the groundstate $^3D_1$ of Sc^+ is 6.54 eV. The baricenter of 3d4s is situated at 0.09 eV above this groundstate, whereas both the baricenter of $3d^2$ (44 of 45 states have been located and the last calculated) at 1.24 eV and the unique state $^1S_0$ of $4s^2$ at 1.45 eV have comparable energies showing the extreme deviation from a linear variation of the configuration energy as a function of the number (2, 1, and 0) of 4s electrons. Equation (7) produces a parabolic variation with this number, and another example of minimum energy for one 4s electron is the neutral nickel atom. Though the groundstate $^3F_4$ belongs to $3d^84s^2$ having the baricenter 1.29 eV higher, the lowest baricenter is that of $3d^94s$ at 0.17 eV, and the unique state $^1S_0$ of $3d^{10}$ is situated at 1.83 eV above the groundstate. The situation is different in the neutral palladium atom with $4d^{10}$ below $4d^95s$, again below $4d^85s^2$.

Using the behavior of a three-electron system such as the scandium atom as a model for transition group compounds, the excitation energy $h\nu$ of electron transfer bands ($3d4s^2 \rightarrow 3d^24s$) and the photoelectron ionization energy I(d) corresponding to ($3d4s^2 \rightarrow 4s^2$) and I(s) ($3d4s^2 \rightarrow 3d4s$) from Eq. (8) can be written

$$\begin{aligned} h\nu &= \mathrm{W(s) - W(d) + J(d, d) - J(d, s)} = 2.03\ \mathrm{eV}, \\ \mathrm{I(d)} &= \mathrm{W(d)} - 2J\mathrm{(d, s)} = 7.98\ \mathrm{eV}, \\ I\mathrm{(s)} &= \mathrm{W(s) - J(d, s) - J(s, s)} = 6.62\ \mathrm{eV}, \end{aligned} \tag{9}$$

with the numerical values taken from Sc and Sc^+. The only reason why $h\nu$ is positive (though W(d) = 24.75 eV in Sc^{+2} is larger than W(s) = 21.60 eV) is the unusually large value of J(d, d). In a certain sense, it is the positive $h\nu$ constituting the surprising fact rather than I(d) being larger than I(s). However, the positive wavenumber of the electron transfer bands has been known for a longer time than the I values in photoelectron spectra.

It is quite obvious that the 4s orbital is not a perfect analogy to the ligand orbitals in the 3d group compounds though the comparison was already made by Orgel (1955) discussing the moderate decrease in the term distance between ^4G and ^6S in the neutral manganese atom $3d^54s^2$, relative to Mn^{+2}, $3d^5$ as a model for the weak nephelauxetic effect in Mn(II) compounds. However, it is clear that situations similar to Eq. (9) occur when $\mathrm{W(d) > W(s)}$ and $\mathrm{J(d, d) \gg J(d, s) \simeq J(s, s)}$. Actually, an approximate equality between J(d, s) and J(s, s) in the solids produces a difference W(d) − W(s) closely similar to I(d) − I(s). The scandium atom is less extreme in this respect since the W difference 3.15 eV is considerably larger than the I difference 1.36 eV (cf. Jørgensen, 1974).

After obtaining a semiquantitative understanding of the simultaneous positive value of $h\nu$ and of $I(\mathrm{d}) - I(\mathrm{s})$ an awesome problem remains which colloquially (Jørgensen, 1973c) is called "the third revolution in ligand field theory." It is already difficult enough for the chemist to define clearly what is meant with one-electron energy. Thus, it is by no means a trivial dilemma that the scandium atom prefers to contain two 4s electrons leaving the third to the 3d shell, though the 4s electrons are most readily lost by ionization. This discrepancy between electron energetic order derived from population in the ground configuration and derived from photoelectron I values remains also in compounds. However, a far more profound difficulty is that the condition for strong covalent bonding usually is thought of as identical diagonal elements of energy. Hence, the results summarized in Table III suggest that ytterbium(III) compounds are strongly covalent in blatant disagreement with the spectroscopic fact that the distance from the excited level $^2F_{5/2}$ to the groundstate $^2F_{7/2}$ invariably is 1.3 eV, showing that the radial 4f function hardly is modified by chemical bonding. The same conclusion is drawn from the weak nephelauxetic effect (Jørgensen, 1962a, 1969a) in M(III) compounds containing from two to twelve 4f electrons. It might be argued that the actual degeneracy allowing an arbitrary linear combination of eigenfunctions to be a solution to Schrödinger's equation occurs between M(IV) having lost a 4f electron and the ligands having lost an electron. There is little doubt that 4f group M(IV) compounds are more covalent than M(III) but still, the nephelauxetic effect does not seem to be excessive in Nd(IV) and Dy(IV) double fluorides, and negative evidence is that transitions to ^6P of $4f^7$ have not been detected at wavenumbers lower than those of the electron transfer bands in terbium(IV) compounds.

If one abandons the hypothesis that the diagonal elements of a Hückel-type model of transition group compounds are related to ionization energies, it is possible to take the critical attitude that since the total energy anyhow is not the sum of the one-shot ionization energies, the total energy may very well decrease at the cost of a few electrons increasing their energy. It is obvious that the variation principle applies only to the total wavefunctions though it becomes less convincing because the numerical value of the correlation energy is larger than the first ionization energy of atoms with Z above 10. Hence, an infinite number of levels having the same symmetry type as the groundstate occurs below the Hartree–Fock energy. Nevertheless, one may cite several reasons to hope for an effective one-electron operator producing orbital energies as eigenvalues. Going from Ce(III) to Yb(III), the 4f ionization energy cross the ligand orbitals at least once, and normally thrice (Wertheim *et al.*, 1971; Jørgensen, 1973a) because of the spin-pairing energy; I(M4f) is always higher than of the ligands for Gd and usually below for Tb. Nevertheless, one does not observe any discontinuities in the "ligand field" separation of the seven 4f energies, and it is beyond doubt that the parameters of the angular overlap model have the same sign (Jørgensen *et al.*, 1963; Burns, 1967) as in d group complexes corresponding to antibonding of the orbitals directed toward the ligands. After all, it would have been conceivable to have the opposite sign expressing back-bonding from the filled ligand orbitals at higher energy.

The probable conclusion is that one-electron energy *differences* within the same *nl*-shell of the central atom are accounting quantities able to describe the observed energy differences between low-lying levels of the complex, but do not *per se* represent differences in ionization energies. A serious problem is then what are the eigenvectors in the LCAO approximation. The highly delocalized and covalent character of Yb(IV) having lost one of the thirteen 4f electrons in a photoelectron experiment does not occur in the 14 states constituting the two ^{2}F levels belonging to the ground configuration $4f^{13}$ of Yb(III) hardly modified by the chemical bonding. One might imagine various solutions of this problem. In highly ionic systems MX, it is perhaps the electron affinity of M and the ionization energy of X which determine the extent of covalent bonding. This argument would need a certain amount of reiteration in order to find whether a starting approximation of more or less covalent bonding is consistent. Another solution of somewhat larger generality (not needing a preliminary look at "the facit list") would be to assume that diagonal elements H_M and H_X of the effective one-electron operator are intended to indicate the Mulliken electronegativity (Jørgensen, 1962a), the arithmetic average of the electron affinity and the ionization energy of the orbital considered. This definition would not change the Hückel treatment of a heterocyclic organic molecule to any large extent because carbon and nitrogen atoms have comparable J(2p, 2p), but the Mulliken electronegativity is strongly decreased (Jørgensen, 1969a) in transition group atoms relative to the I values. The question why ytterbium(III) is much less covalent than copper(II) compounds having similar Mulliken electronegativity of the central atom can then be answered by reference to the much smaller overlap integral S_{MX} in the former case. It is quite plausible, at least in the limiting case of weak covalent bonding, that the nondiagonal element of the effective one-electron operator is indeed proportional to S_{MX}. However, one might deviate from the Wolfsberg–Helmholz model by suggesting (Jørgensen, 1967, 1971b) that the energetic multiplier is not a suitably chosen average of H_M and H_X but the asymptotic value of the *kinetic energy operator* in the bond region of the two atomic orbitals forming the chemical bond. This suggestion is compatible with all the results of the angular overlap model known to represent the d group internal transitions in partly filled shells completely. On the other hand, it was not realized in 1970 that the diagonal elements H_M and H_X are perhaps not exactly ionization energies. Quite generally, ligand field theorists tend to spend an enormous effort on details and do not think much about the more fundamental problems behind the approximative models.

In our study (Jørgensen and Berthou, 1972a) of photoelectron spectra of solids, we find the 3d signals of Cr(III) close to $I = 10$ eV (cf. Helmer, 1973) as is also true for both Co(II) and Co(III). I(3d) for nickel(II) is distributed from 12 eV in the least covalent to 9 eV in the most covalent compounds. Though I'(3d) is only 9.2 eV for $KNiF_3$, a chemical shift of I(3d) has also been reported (Evans *et al.*, 1973) in gaseous chromium(III) acetylacetonate at 7.5 eV, that is, 2.1 eV below the hexafluoroacetylacetonate. Whereas most of spin-allowed transitions studied in ligand field theory (Griffith, 1961; Jørgensen, 1962b) involve antibonding d-like electrons, there is only a minority of complexes containing such antibonding electrons in the groundstate. However, the upper subshell of octahedral Mn(II), Fe(II), Co(II), and Ni(II) complexes contain two electrons, and one might hope for separate signals from the two subshells though they may not show exactly the same energy difference (1–1.5 eV) as derived from the absorption spectra because of the effects (Jørgensen, 1958) of interelectronic repulsion discussed above. Actually, the 3d signal of cobalt(II) and nickel(II) fluorides and hexaqua salts show a shoulder about 1 eV below I of the peak presumably due to the two and five (or six) d-like electrons, respectively. However, an additional complication may occur because of the accessibility of several quintet and triplet levels of $3d^6$ and quartet and doublet levels of $3d^7$ formed. This subject awaits measurements of better resolution. I(3d) is not much larger for the copper(II) fluoride (13 eV) than copper(I) cyanide (11 eV). The two I' values are both 2 eV lower. A major difference from the 4f group (where I(4f) is marginally larger for Gd(III) than for Lu(III) salts of the same anion, and larger for TbO_2 than for HfO_2) is that I(3d) is close to 18 eV in zinc(II) and 28 eV in gallium(III) compounds, much higher than for systems in the same oxidation state containing partly filled 3d shells. I(4d) values between 9 and 11 eV are observed (Jørgensen and Berthou, 1972a) for many rhodium(III) and palladium(II) compounds and 11 to 13 eV in silver(I) as contrasted to about 19 eV in cadmium(II) and 26 eV in indium(III) compounds. On the whole, I(5d) values are slightly lower. It is seen that for partly filled shells, the d ionization energies are most frequently below the ligand orbitals as presumed in older ligand field theory and known for gaseous VCl_4 and $M(CO)_6$ having $I = 8.5$ eV for M = Cr, Mo, and W (Turner *et al.* 1970).

It is also possible to compare with metallic elements (Fadley and Shirley, 1970b; Baer *et al.*, 1970) where silver has a signal at $I^* = 4.7$ eV with a shoulder at 5.5 eV and gold two distinct signals at $I^* = 3.2$ and 6.0 eV. The I^* values are relative to the Fermi level and should be increased by the work function (close to 4.7 eV) in order to represent I values relative to *vacuo*. There is no doubt that the separation 2.8 eV of the two Au5d signals is not due exclusively to spin-orbit coupling (of which the effect is only 1.5 eV in mercury(II) compounds) and it has been argued that the shape of the photoelectron spectrum indicates the density of states in the energy band model assuming roughly constant probabilities of ionization by the soft X rays. On the other hand, alloys of palladium and silver (Hedman *et al.*, 1971) show distinct Pd4d signals at 3 eV lower I^* than the Ag4d signals as if the orbitals are localized.

Metallic lanthanides fall in two groups (Jørgensen, 1969a) according to whether the number of 4f electrons corresponds to the conditional oxidation state M[II] as is true for the barium-like europium and ytterbium, or M[III] as the rest. For comparison with Table III, it may be noted that metallic ytterbium (Hagström *et al.*, 1971) has $I^*(4f) = 1.4$ and 2.7 eV. The reason why the lowest I (relative to *vacuo*) close to 4 eV is so much lower than for Yb(III) is the barium-like character of the metal. On the other hand, metallic gadolinium has $I^* = 7.8$ eV and $I = 11.0$ eV to be compared with $I'(4f) = 15.3$ eV for GdF_3 and 13.6 eV for Gd_2O_3 (Jørgensen and Berthou, 1972a). The moderate shift from Gd(III) compounds to metallic Gd[III] is quite typical for other elements. Again, it is an intriguing problem how the energy band model can be brought in agreement with a metallic lanthanide showing a highly positive I^* for the partly filled 4f shell. The situation of filled MO's having lower I than the partly filled shell in nonmetallic compounds discussed above is analogous to the conduction electrons of the metal not invading the partly filled 4f shell which has an electron affinity considerably smaller than the Fermi level though its ionization energy is the other way around (Cox *et al.*, 1973).

Recent MO calculations of transition group complexes showing strong deviations from Koopmans behavior include a study by Demuynck and

Veillard (1973) of $Ni(CO)_4$ where the calculated ε_i for the loosest bound MO's of symmetry types $9t_2$, $2e$, and $8t_2$ are 10.75, 12.82, and 17.52 eV, to be compared with the experimental "vertical" ionization energies 8.8, 9.7, and 14.8 eV showing the subshell energy difference 0.9 eV in this $3d^{10}$ system. The situation is far more confusing in the diamagnetic ($S = 0$) $3d^8$ complex $Ni(CN)_4^{-2}$ having all nine nuclei on a Greek cross and calculated to have not less than eight MO's mainly delocalized on the cyanide ligands with ε_i between 3.29 and 4.35 eV above the first mainly $3d(3z^2 - r^2)$-like orbital at 4.95 eV. Two other $3d(xz, yz)$ occur at 7.26 and the last (xy) at 9.1 eV. Nevertheless, the first spin-allowed transition $(3z^2 - r^2) \rightarrow (x^2 - y^2)$ is calculated at 2.56 eV in fair agreement with the observed shoulder at 3.4 eV. Crystalline $K_2Ni(CN)_4$ has $I(\text{Ni3d}) = 9$ eV. The chemical difficulty of oxidizing $Ni(CN)_4^{-2}$ in aqueous solution suggests an adiabatic ionization energy of the loosest bound MO's above 6 eV and the "vertical" I even higher.

Novakov (1971) detected rather unusual satellites accompanying the $Cu2p_{1/2}$ and $Cu2p_{3/2}$ signals of many copper compounds. It has now been firmly established (Jørgensen and Berthou, 1972a; Frost *et al.*, 1972b) that these satellites only occur in Cu(II) and not in the closed-shell Cu(I). In view of the oxidizing character of Cu(II) frequently showing up as electron transfer bands in the visible, one might have expected the satellites to be due to electron transfer to the vacant $3d(x^2 - y^2)$-like position in analogy to La(III) stabilizing its empty 4f shell by inner-shell ionization. However, this is not too probable in view of the fact (Frost *et al.*, 1972b) that the satellite distance is 7 eV with the least reducing ligands, fluoride, and slowly increases to above 9 eV when $I(\text{Cu2p})$ decreases in the more covalent compounds such as the oxide and the acetylacetonate. The Cu3d signal seems to have a comparable satellite. When the $I(\text{Cu3d})$ values given above are considered, it is not possible to suggest the simultaneous ejection of two 3d electrons, in particular because it needs more than twice I for one electron according to Eq. (7). Hence, the most plausible explanation is intra-atomic shake-up of the 3d shell. It is not easy to see why exactly copper(II) should show such satellites more than many other central atoms, and in particular, it needs less energy to excite the 3d electrons of copper(I) to empty shells such as 4s or 4p participating in the formation of antibonding MO's. Some authors argue that only paramagnetic (positive S) central atoms show these shake-up satellites *strongly*. The problem with photoelectron spectra of solids is that the background is very intense and mildly undulating, making the detection of weak satellites very difficult. However, there is one rather unique property of copper(II), the single antibonding electron in a $(x^2 - y^2)$-type orbital strongly delocalized on the ligating atoms. The probability of shake-up may be enhanced by this delocalization. Along this line of thought, some authors then would argue that high-spin Mn(II), Fe(II), Co(II), and Ni(II) also contain antibonding electrons, though less delocalized because of less pronounced covalent bonding, and what we take as $(S \pm \frac{1}{2})$ structure due to differing interelectronic repulsion is indeed satellites of the same kind as in Cu(II). One difficulty is that diamagnetic quadratic Ni(II) and Pd(II) complexes do not show satellites, and that the weak satellites in quadratic Cu(III) most probably are due to partial superficial reduction to Cu(II). It is usually argued (Krause *et al.*, 1968) that the orbital selection rule for shake-up is that of monopole transitions, conserving the symmetry type in the prevailing point group. In the spherical symmetry of monoatomic entities, this means invariant l. If the monopole selection rule was important, the totally symmetric $(3z^2 - r^2)$ orbital in quadratic complexes could shake up to 4s whereas the opportunities in octahedral symmetry would be far more restricted. However, Hamnett (1973) express doubts whether this selection rule is realistic, and points out that only the $M(C_5H_5)_2$ containing two antibonding d-like electrons (M = Mn and Ni) show extensive satellite structure in the M2p region. The problem of these specific Cu(II) type satellites is most fascinating and by no means resolved.

VI. The Question of Time-Scale and the Manne Effect

It is perfectly consistent that physicists looking for a different time a the same system sometimes have entirely different ideas about the symmetry and other properties. At one extreme, the diffraction by a crystal of X rays or neutrons produce a time-average picture of the average content of the unit cell. At the other extreme, it is a legitimate question in quantum mechanics how a molecule looks on an *instantaneous picture*. In the following, we make the simplification of considering the nuclei as points. It is quite evident that smaller molecules, such as methane or copper(II) phthalocyanine, exist as reproducible entities. We do not here enter a discussion with Professor Primas about the subtle question whether a sufficiently large molecule, or a cat, is so unique that it no longer is a carrier of assembly properties. Anyhow, it is perfectly possible to device a Gedanken experiment where the positions of all five nuclei are found in a methane molecule. Because of the vibration (even at 0°K) these positions do not normally form a regular tetrahedron with its center; only the average of a large number of such measurements is situated in a statistical sense at such a symmetrical structure. Actually, the instantaneous point group is almost always the lowest possible C_1 in the same strong sense as the probability of selecting an irrational number among the set of real numbers is closer to 100% than any positive quantity. This argument does not hold for triatomic molecules not getting below the point group C_s because three points always define a plane, nor for heteroatomic and homoatomic dinuclear molecules which cannot lose their symmetry $C_{\infty v}$ and $D_{\infty h}$ by vibration. Anyhow, the methane molecule possesses the point group T_d in the sense that the selection rules for vibrational spectra obtained by infrared or Raman techniques are those applicable to the average positions of the nuclei, and what is more important for our purposes, no separations are observed between two or three degenerate normal frequencies coinciding for group-theoretical reasons.

Some molecules refuse systematically the highest possible symmetry available to the nuclei (Jørgensen, 1969d, 1971b). This situation occurs for Gillespie unstable systems (Gillespie, 1972) and for first-order Jahn–Teller unstable systems. In the approximation of a well-defined subshell configuration, those d group complexes are Jahn–Teller unstable (Jørgensen, 1962b) which contains a differing number (0 and 1; or 1 and 2) of electrons in two or three orbitals having the same energy of group-theoretical necessity. Pearson (1969) prefers to call Gillespie unstable systems for second-order Jahn–Teller instability. However, besides the historical reasons that the optical activity of sulphonium cations $RR'R''S^+$ with three different substituents was described by Sidgwick as tetrahedral species with one lone pair and three other ligands, the point is that Gillespie instability is a numerical question of the energetic advantage of *l*-mixing relative to the restoring forces preventing the distortion. It may be noted that orbitals with well-defined *l* have an electronic density of even parity, whereas a mixture of even and odd orbitals is not obliged to have well-defined parity of its square. The Gillespie description (Gillespie, 1972) is normally applied to p group complexes such as square-pyramidal TeF_5^-, IF_5, and XeF_5^+ being octahedra with the sixth ligand replaced by a lone pair, but some residual traces are found at the end of the d groups, such as the unusual stereochemistry and high absorption band intensities in copper(II) complexes (Jørgensen, 1963b, 1971b) and complex formation constants and intensities in palladium(II) complexes (Rasmussen and Jørgensen, 1968).

One of the simplest Gillespie unstable molecules is ammonia. At room temperature and below, where the excited vibrational states are not populated significantly according to the Boltzmann formula, the umbrella inversion frequency going from one pyramidal form to the opposite is about 10^{10} per sec. At higher temperatures, a diffuse superposition of vibrational wavefunctions would occur with a pronounced minimum for the planar configuration belonging to the point group D_{3h} like BF_3, but there is only a difference of degree and not of essence between this case and crystal structures of apparent high symmetry showing anomalous anisotropy of the thermal vibration. Other examples of Gillespie unstable species becoming pyramidal are SO_3^{-2}, ClO_3^-, $SnCl_3^-$, $SbCl_3$, $TeCl_3^+$, IO_3^-, and XeO_3.

From a quantum-mechanical point of view, the conventional Born–Oppenheimer factorization of translational, rotational, vibrational, and electronic wavefunctions induces a continuous spectrum of translational energy starting at the groundstate of any system which is not heavily confined in a small volume. Nevertheless, the spectral lines of a neon sign placed on the roof of a tram are not perceptibly broadened, though they show a minute Doppler shift. This absence of coupling between continuous and discrete quantum numbers is also seen in the narrow photoelectron signals and in X-ray absorption edges where the excited states technically are situated in the continuum of electronic states but show singularities in other physical properties (such as transition probabilities) because of the validity of the one-electron classification (Jørgensen, 1969a, 1971b).

Another, related, question is how long-lived states need to be before they are considered stationary. The optically active enantiomers of a molecule cannot be eigenfunctions of the Schrödinger equation though the racemization may take millions of years. By the same token, the umbrella inversion of AsH_3 is calculated to take many months. It may be noted that disregarding translational and rotational motion, the complete set of internuclear distances in a polyatomic species is sufficient to describe it without further reference to the point-group symmetry, with the explicit exception of optical activity.

Since the wavenumber 3300 cm^{-1} corresponds to a characteristic time of 10^{-14} sec and 330 cm^{-1} to 10^{-13} sec, it is generally argued that visible and ultraviolet spectra obeying Franck–Condon's principle produce a picture of the absorbing species within this time limit. Frequently, the importance of the *time scale* of a given experiment can be determined. Thus, by varying the temperature of a crystal of $[Cu(H_2O)_6]SiF_6$, the electron paramagnetic resonance spectrum shows an isotropic g value corresponding to an average picture of a regular octahedron above 100°K whereas a g tensor with a third value deviating strongly from the two others can be seen at lower temperatures. Like in comparable cases of nuclear magnetic resonance, the influence of the temperature is both a change of the time scales of the measurement and of the rearrangement of the internuclear distance. The Jahn–Teller unstable $Cu(H_2O)_6^{+2}$ both in this crystal and in aqueous solution has a spectrum in the near infrared (Jørgensen, 1963b) suggesting four short Cu-O distances along two of the Cartesian axes and two long distances on the perpendicular axis. The reorientation of the elongated direction takes about 10^{-11} sec in both water and methanol at 25°C (Poupko and Luz, 1972). This time should not be confused with the time about 10^{-8} sec needed to exchange a ligated molecule in the first coordination sphere. It may be mentioned that both $Cu(NH_3)_5^{+2}$ and the blue tetrammine in aqueous solution which has turned out (Romano and Bjerrum, 1970) to be $Cu(NH_3)_4(H_2O)^{+2}$ are square-pyramidal like XeF_5^+ whereas the anhydrous, truly quadratic $Cu(NH_3)_4^{+2}$ is pink (Jørgensen, 1971b).

One may ask whether the picture obtained by photoelectron spectra corresponds to an even shorter time scale than visible spectra. It was a dogmatic truth in the Copenhagen version of quantum mechanics that quantum jumps do not take any time at all. This postulate is a non-relativistic statement. It takes light or X rays 10^{-18} sec to cross an atom having the diameter 3 Å, and it would be difficult to believe in a shorter induction time of the primary ionization process. If the time of passage of one wavelength of the X rays is taken as a reference, 3.10^{-18} sec would be obtained. One can make several arguments why the primary ionization process takes less than 10^{-17} sec. For instance, the broadest signal we have identified (4p of cadmium) corresponds to a half-life 5.10^{-17} sec according to Heisenberg's uncertainty principle. Hence, only a few layers of neighbor atoms can be contributing to the decreased I by relaxation (excepting pre-existing polarization of the groundstate) if one assumes that the information about the primary ionization has to go forward to the relaxing atom and back again within the time of the process, and that it is not possible locally to extract information about the future. It is well known from nuclear reactions how particles moving with a thirtieth of the speed of light cross 10^{-12} cm during 10^{-21} sec. However, the situation is different because the nuclei normally are perfectly individual entities, and there is only one of each kind in a large volume, whereas the electrons are indiscernible, and occur in a cloud close to the nucleus. Nevertheless, it is reasonable to say that the X-ray-induced photoelectron spectra is the most rapid technique available for studying electronic structure at present, and more rapid than visible spectra.

One aspect of the characteristic time scale of order of magnitude 10^{-17} sec is that electronic ordering (Jørgensen, 1969a; Robin and Day,

1967) of mixed-valence compounds is frequently (or always) observed. Also, it is not surprising that the hydrogen bond in protonated 1,8-bis-(dimethylamino)naphtalene (Haselbach *et al.*, 1972) seems slightly asymmetrical, two N1s signals are observed at 1.4 eV distance, somewhat smaller than the usual difference between protonated and neutral amino groups. A comparable problem (Clark *et al.*, 1971) involves the thiathiophtens studied by Klingsberg having two long, but differently long, sulfur–sulfur bonds.

Relaxation effects in the electronic density of neighbor atoms can be important, not only in nonmetallic and metallic solids, but also in polyatomic molecules in the gaseous state. Perry and Jolly (1972) argue that the values shown in Table V indicate bonding between the ligand π orbitals and the empty (rather high) silicon 3d and germanium 4d with exception of the purely σ-bonded hydrides. However, an alternative explanation is that Manne effect is weaker in the hydride than in the case of the Si and Ge compounds of the heavier ligands. The applicability of a linear model (Stucky *et al.*, 1972) including the Madelung potential from the fractional charges on adjacent atoms has been critically evaluated. Assuming atomic charges from *ab initio* calculations (which seem on the low side, for instance +1.0 on carbon and −0.25 on fluorine in CF_4), the slopes 20.7 eV per charge on C1s and 31 eV for the N1s signal are obtained. Apparently, the reorganization energy can be incorporated to a large extent in this potential model, though reservations are made (Stucky *et al.*, 1972) regarding strongly conjugated systems. The estimate for a neutral carbon atom in a gaseous molecule is 291.7 eV whereas I relative to *vacuo* for graphite (Shirley, 1972) is close to 288 eV.

TABLE V

I (eV)

M =	C1s	$Si2p_{3/2}$	$Ge3p_{3/2}$
$M(CH_3)_4$	290.3	105.8	120.6
MH_4	290.7	107.1	123.7
MCl_4	296.2	110.2	124.2
MF_4	301.7	111.5	128.9

Though ESCA means "electron spectroscopy for chemical analysis" the writer is convinced that the help toward understanding of chemical bonding has been far more important than the analytical applications. In particular, the experimental determination of I of penultimate orbitals has been interesting and confirms the utility of the one-electron description and the classification of molecular orbitals. Further on, the old question of I(4f) of the rare earths (Jørgensen, 1973a) has been answered, and at the same time, the semiempirical estimate of covalent bonding involving partly filled shells has met a highly unexpected challenge showing how the behavior of the transition groups can only be understood if interelectronic repulsion is carefully considered. Finally, the questions of the time scale of a given experimental technique strongly influence the opinion formed about the symmetry of a molecule or polyatomic ion. Quite generally, the continuum contains singularities producing photoelectron signals with half-widths determined partly by the Franck–Condon principle based on unchanged internuclear distances and partly by the uncertainty principle of Heisenberg. The subsequent emission of Auger electrons is comparable to the phenomenon of autoionization of high-lying energy levels in atomic spectroscopy. Nevertheless, the progress in this field has been so recent that one can hope for still many more surprises.

ACKNOWLEDGMENTS

The writer is grateful to the Swiss National Science Foundation for the grant to purchase the photoelectron spectrometer, and to Dr. Hervé Berthou and Dr. Lucette Balsenc for their careful and indefatigable experimental effort.

REFERENCES

AARONS, L. J., GUEST, M. F., HALL, M. B., and HILLIER, I. M. (1973). *Trans. Faraday Soc.* **69**, part II, p. 563.

AHRLAND, S. (1968). *Struct. Bonding* (*Berlin*) **5**, 118.

BAER, Y., HEDÉN, P. F., HEDMAN, J., KLASSON, M., NORDLING, C., and SIEGBAHN, K. (1970). *Phys. Scripta* (*Stockholm*) **1**, 55.

BAMBYNEK, W., CRASEMANN, B., FINK, R. W., FREUND, H. U., MARK, H., SWIFT, C. D., PRICE, R. E., and VENUGOPALA RAO, P. (1972). *Rev. Mod. Phys.* **44**, 716.

BASCH, H., VISTE, A., and GRAY, H. B. (1966). *J. Chem. Phys.* **44**, 10.

BONNELLE, C., and JØRGENSEN, C. K. (1964). *J. Chim. Phys. Physicochim. Biol.* **61**, 826.

BONNELLE, C., KARNATAK, R. C., and JØRGENSEN, C. K. (1972). *Chem. Phys. Lett.* **14**, 145.

BREMSER, W., and LINNEMANN, F. (1971). *Chem.-Ztg., Chem. App.* **95**, 1011.

BRUNDLE, C. R., ROBIN, M. B., and BASCH, H. (1970). *J. Chem. Phys.* **53**, 2196.

BURNS, G. (1967). *Phys. Lett. A.* **25**, 15.

CITRIN, P. H., and THOMAS, T. D. (1972). *J. Chem. Phys.* **57**, 4446.

CITRIN, H. and HAMANN, D. R. (1973). *Chem. Phys. Letters* **22**, 301.

CITRIN, P. H., HAMANN, D. R., and HÜFNER, S. (1973). *Bull. Amer. Phys. Soc.* **18**, 55.

CLARK, D. T., KILCAST, D., and REID, D. H. (1971). *Chem. Commun.* p. 638.

COHEN, R. L., WERTHEIM, G. K., ROSENCWAIG, A., and GUGGENHEIM, H. J. (1972). *Phys. Rev. B* **5**, 1037.

CONDON, E. U., and SHORTLEY, G. H. (1953). "Theory of Atomic Spectra." Cambridge Univ. Press, London and New York.

CONNOLLY, J. W. D., SIEGBAHN, H., GELIUS, U., and NORDLING, C. (1973). *J. Chem. Phys.* **58**, 4265.

COX, P. A., EVANS, S., and ORCHARD, A. F. (1970). *Chem. Phys. Lett.* **7**, 414.

COX, P. A., BAER, Y., and JØRGENSEN, C. K. (1973). *Chem. Phys. Lett.* **22**, 433.

DEMUYNCK, J., and VEILLARD, A. (1973). *Theor. Chim. Acta* **28**, 241.

EVANS, S., HAMNETT, A., and ORCHARD, A. F. (1972). *J. Coord. Chem.* **2**, 57.

EVANS, S., HAMNETT, A., ORCHARD, A. F., and LLOYD, D. R. (1973). *Discuss. Faraday Soc.* **54**, 227.

FADLEY, C. S., and SHIRLEY, D. A. (1970a). *Phys. Rev. A* **2**, 1109.

FADLEY, C. S., and SHIRLEY, D. A. (1970b). *J. Res. Nat. Bur. Stand., Sect. A* **74**, 543.

FADLEY, C. S., HAGSTRÖM, S. B. M., KLEIN, M. P., and SHIRLEY, D. A. (1968). *J. Chem. Phys.* **48**, 3779.

FADLEY, C. S., SHIRLEY, D. A., FREEMAN, A. J., BAGUS, P. S., and MALLOW, J. V. (1969). *Phys. Rev. Lett.* **23**, 1397.

FAESSLER, A. (1972). *Angew. Chem.* **84**, 51.

FENSKE, R. F., CAULTON, K. G., RADTKE, D. D., and SWEENEY, C. C. (1966). *Inorg. Chem.* **5**, 951, 960.

FINN, P., PEARSON, R. K., HOLLANDER, J. M., and JOLLY, W. L. (1971). *Inorg. Chem.* **10**, 378.

FROST, D. C., MCDOWELL, C. A., and WOOLSEY, I. S. (1972a). *Chem. Phys. Lett.* **17**, 320.

FROST, D. C., ISHITANI, A., and MCDOWELL, C. A. (1972b). *Mol. Phys.* **24**, 861.

GASPAR, R. (1967). *Int. J. Quantum Chem.* **1**, 139.

GILLESPIE, R. J. (1972). "Molecular Geometry." Van Nostrand-Reinhold, Princeton, New Jersey.

GRIFFITH, J. S. (1961). "Theory of Transition-Metal Ions." Cambridge Univ. Press, London and New York.

HAGSTRÖM, S. B. M., BRODÉN, G., HEDÉN, P. O., and LÖFGREN, H. (1971). *J. Phys. (Paris) C* **4**, 269.

HAMNETT, A. (1973). *Discuss. Faraday Soc.* **54**, 251.

HASELBACH, E., HENRIKSSON, A., JACHIMOWICZ, F., and WIRZ, J. (1972). *Helv. Chim. Acta* **55**, 1757.

HEDMAN, J., KLASSON, M., NILSSON, R., NORDLING, C., SOROKINA, M. F., KLJUSHNIKOV, O. I., NEMNONOV, S. A., TRAPEZNIKOV, V. A., and ZYRYANOV, V. G. (1971). *Phys. Scripta (Stockholm)* **4**, 1.

HELMER, J. C. (1973). *J. Electron Spectrosc.* **1**, 259.

HENDRICKSON, D. N., HOLLANDER, J. M., and JOLLY, W. L. (1969). *Inorg. Chem.* **8**, 2642.

HERZBERG, G. (1966). "Electronic Spectra and Electronic Structure of Polyatomic Molecules." Van Nostrand-Reinhold, Princeton, New Jersey.

HNATOWICH, D. J., HUDIS, J., PERLMAN, M. L., and RAGAINI, R. C. (1971). *J. Appl. Phys.* **42**, 4883.

JØRGENSEN, C. K. (1958). *Acta. Chem. Scand.* **12**, 903.

JØRGENSEN, C. K. (1962a). "Orbitals in Atoms and Molecules." Academic Press. London and New York.

JØRGENSEN, C. K. (1962b). "Absorption Spectra and Chemical Bonding in Complexes." Pergamon, Oxford.

JØRGENSEN, C. K. (1963a). "Inorganic Complexes." Academic Press, New York.

JØRGENSEN, C. K. (1963b). *Advan. Chem. Phys.* **5**, 33.

JØRGENSEN, C. K. (1967). *Chem. Phys. Lett.* **1**, 11.

JØRGENSEN, C. K. (1968). *Bull. Soc. Chim. Fr.* p. 4745.

JØRGENSEN, C. K. (1969a). "Oxidation Numbers and Oxidation States." Springer-Verlag, Berlin and New York.

JØRGENSEN, C. K. (1969b). *Chimia* **23**, 292.

JØRGENSEN, C. K. (1969c). *Rev. Chim. Miner.* **6**, 183.

JØRGENSEN, C. K. (1969d). *Chem. Phys. Lett.* **3**, 380.

JØRGENSEN, C. K. (1970). *Progr. Inorg. Chem.* **12**, 101.

JØRGENSEN, C. K. (1971a). *Chimia* **25**, 213.

JØRGENSEN, C. K. (1971b). "Modern Aspects of Ligand Field Theory." North-Holland Publ., Amsterdam.

JØRGENSEN, C. K. (1972). *Theor. Chim. Acta* **24**, 241.

JØRGENSEN, C. K. (1973a). *Struct. Bonding (Berlin)* **13**, 199.

JØRGENSEN, C. K. (1973b). *Angew. Chem.* **85**, 1.

JØRGENSEN, C. K. (1973c). *Chimia* **27**, 203.

JØRGENSEN, C. K. (1974), *Chimia* **28**, 6.

JØRGENSEN, C. K., and BERTHOU, H. (1972a). *Kgl. Dan. Vidensk. Selsk., Mat. Fys. Medd* **38**, No. 15.

JØRGENSEN, C. K., and BERTHOU, H. (1972b). *Chem. Phys. Lett.* **13**, 186.

JØRGENSEN, C. K., and BERTHOU, H. (1973). *Discuss. Faraday Soc.* **54**, 269.

JØRGENSEN, C. K., PAPPALARDO, R., and SCHMIDTKE, H. H. (1963). *J. Chem. Phys.* **39**, 1422.

JØRGENSEN, C. K., HORNER, S. M., HATFIELD, W. E., and TYREE, S. Y. (1967). *Int. J. Quantum Chem.* **1**, 191.

JØRGENSEN, C. K., BERTHOU, H., and BALSENC, L. (1972). *J. Fluorine Chem.* **1**, 327.

JOLLY, W. L. (1973). *Discuss. Faraday Soc.* **54**, 13.

KOOPMANS, T. (1933). *Physica (Utrecht)* **1**, 104.

KRAUSE, M. O., CARLSON, T. A., and DISMUKES, R. D. (1968). *Phys. Rev.* **170**, 37.

LINDHOLM, E., FRIDH, C., and ÅSBRINK, L. (1973). *Discuss. Faraday Soc.* **54**, 127.

LÖWDIN, P. O., and YOSHIZUMI, H. (1959). *Advan. Chem. Phys.* **2**, 207, 323.

MANNE, R., and ÅBERG, T. (1970). *Chem. Phys. Lett.* **7**, 282.

MURRELL, J. N., and RALSTON, B. J. (1972). *Trans. Faraday Soc.* **68**, part II, p. 1393.

NOVAKOV, T. (1971). *Phys. Rev. B* **3**, 2693.

NUGENT, L. J., BAYBARZ, R. D., BURNETT, J. L., and RYAN, J. L. (1971). *J. Inorg. Nucl. Chem.* **33**, 2503.

ORGEL, L. E. (1955). *J. Chem. Phys.* **23**, 1824.

PEARSON, R. G. (1963). *J. Amer. Chem. Soc.* **85**, 3533.

PEARSON, R. G. (1966). *Science* **151**, 172.

PEARSON, R. G. (1968). *J. Chem. Educ.* **45**, 581, 643.

PEARSON, R. G. (1969). *J. Amer. Chem. Soc.* **91**, 4947.

PERRY, W. B., and JOLLY, W. L. (1972). *Chem. Phys. Lett.* **17**, 611.

PIGNATARO, S. (1972). *Z. Naturforsch. A* **27**, 816.

POTTS, A. W., PRICE, W. C., STREETS, D. G., and WILLIAMS, T. A. (1973). *Discuss. Faraday Soc.* **54**, 168.

Poupko, R., and Luz, Z. (1972). *J. Chem. Phys.* **57**, 3311.

Rasmussen, L., and Jørgensen, C. K. (1968). *Acta Chem. Scand.* **22**, 2313.

Robin, M. B., and Day, P. (1967). *Advan. Inorg. Radiochem.* **10**, 248.

Romano, V., and Bjerrum, J. (1970). *Acta Chem. Scand.* **24**, 1551.

Salzmann, J. J., and Jørgensen, C. K. (1968). *Helv. Chim. Acta* **51**, 1276.

Schäffer, C. E. (1973). *Struct. Bonding (Berlin)* **14**, 69.

Schäffer, C. E., and C. K. Jørgensen, (1965). *Mol. Phys.* **9**, 401.

Schwartz, M. E. (1970a). *Chem. Phys. Lett.* **5**, 50.

Schwartz, M. E. (1970b). *Chem. Phys. Lett.* **7**, 78.

Shirley, D. A. (1972). *Chem. Phys. Lett.* **16**, 220.

Siegbahn, K., Nordling, C., Fahlman, A., Nordberg, R., Hamrin, K., Hedman, J., Johansson, G., Bergmark, T., Karlsson, S. E., Lindgren, I., and Lindgren, B. (1967). "ESCA-Atomic, Molecular and Solid-State Structures Studied by Means of Electron Spectroscopy." Almqvist & Wiksell, Stockholm.

Siegbahn, K., Nordling, C., Johansson, G., Hedman, J., Hedén, P. F., Hamrin, K. Gelius, U., Bergmark, T., Werme, L. O., Manne, R., and Baer, Y. (1969). "ESCA Applied to Free Molecules." North-Holland Publ., Amsterdam.

Signorelli, A. J. and Hayes, R. G. (1973). *Phys. Rev. B* **8**, 81.

Slater, J. C., and Johnson, K. H. (1972). *Phys. Rev. B* **5**, 844.

Stucky, G. D., Matthews, D. A., Hedman, J., Klasson, M., and Nordling, C. *J. Amer. Chem. Soc.* **94**, 8009. (1972).

Sugar, J. (1972a). *Phys. Rev. A* **6**, 1764.

Sugar, J. (1972b). *Phys. Rev. B* **5**, 1785.

Turner, D. W., Baker, C., Baker, A. D., and Brundle, C. R. (1970). "Molecular Photoelectron Spectroscopy." Wiley (Interscience), New York.

Verhaegen, G., Berger, J. J., Desclaux, J. P., and Moser, C. M. (1971). *Chem. Phys. Lett.* **9**, 479.

Wagner, C. D. (1972). *Anal. Chem.* **44**, 1050.

Watanabe, K. (1957). *J. Chem. Phys.* **26**, 542.

Wertheim, G. K., and Rosencwaig, A. (1971). *Phys. Rev. Lett.* **26**, 1179.

Wertheim, G. K., Rosencwaig, A., Cohen, R. L., and Guggenheim, H. J. (1971). *Phys. Rev. Lett.* **27**, 505.

Wybourne, B. G. (1965). "Spectroscopic Properties of Rare Earths." Wiley (Interscience), New York.

Some Recent Developments in the Theory of Coordination Compounds of Metals

G. BERTHIER
Laboratoire de Chimie Quantique
Institut de Biologie Physico-Chimique
Paris, France

I. The Concept of Dative Bonds

The coordination compounds of organic chemistry (amine–oxides, amine–boranes) as well as those of inorganic chemistry (transition metal complexes) are a category of well-defined molecules or ions baffling the conventional definitions of the theoretical chemistry. In addition to the two classes of chemical bonds, the first one for the "ionic" compounds and the other for the "covalent" compounds, it seemed to be necessary to coin a third sort of bond for those compounds, the so-called "*semipolar*" or "*dative*" bond and to imagine a special symbol for it, namely $D \rightarrow A$

or even $D^+ \rightarrow A^-$. The arrow means that the electrons forming the chemical bond between D and A come both from the same atom of the donor group D and are going to fill a hole in the valence shell of an atom of the acceptor group A, implying that there should be an intermolecular charge transfer from the donor to the acceptor, the consequence of which is an important bond dipole moment, say $5D$.

From a structural point of view, it is essential to know the extent of the electron transfer in dative bonds. This problem has been the subject of many theoretical works in the last decade, the charge migration being estimated either from classical electrostatic theories or from approximate quantum-mechanical methods. In the seventies, the first *ab initio* calculations on coordination compounds have been achieved, opening a new field to the applications of the nonempirical methods of quantum chemistry. Therefore, it may be opportune to review the present state of our knowledge concerning the nature of the dative bond, just before the subject becomes an almost entirely computational matter.

II. Classical Descriptions of the Electronic Structure of Coordination Compounds

A. The Sidgwick Rule and Its Amendments

We will not give an exhaustive review of the theories which were formulated after the historic work of Werner on coordination compounds, but we wish to simply recall the salient features of the classical theories, because some of their conclusions are still debated in the most recent quantum-mechanical calculations, for instance the splitting of the *d* orbitals of a metal by ligands and its consequence for the atomic net charges through the so-called " back-donation " effect.

The first operative theory concerning the electronic structure of coordination compounds was probably given by Sidgwick (1923) in the form of an extension of the octet rule: The coordination number of an atom in a given oxidation state cannot exceed a fixed value, which is reached when the number of its own electrons plus the total number of bond electrons coming from the ligands reproduces the electronic configuration of the next noble gas in the periodical table. The case of the following complexes of iron is most typical:

$$[\mathrm{Fe(CN)_6}]^{4-}: \mathrm{Z(Fe^{2+})} + 6\ \mathrm{(CN)\ electron\ pairs} = 36,$$
$$\mathrm{Fe(CO)_5}: \mathrm{Z\ (Fe^0)} + 5\ \mathrm{(CO)\ electron\ pairs} = 36.$$

Although the Sidgwick rule may be illustrated by many other examples its justification rests on an analogy with closed-shell atoms, which is not very convincing either from the theoretical point of view or from the experimental one; consider the fact that noble gases are, indeed, able to form chemical compounds. However, this rule cannot be ignored completely, because it has two important consequences for the structure of coordination compounds. First of all, it gives a way to connect the geometrical structure of complexes to their electronic structure (Sidgwick and Powell, 1940) provided that the stereochemistry of an atom in a molecule is assumed to be controlled by the electron pairs of the bonds directed toward its nearest neighbors (including the ones which come from donor atoms of ligands): According to Gillespie and Nyholm (1957) the most stable conformation should correspond to the repulsion maximum between the electron pairs of the effective valence shell of the central atom, that is to say the octahedron for an atom bonded to six identical ligands and so forth. We will discuss the Gillespie–Nyholm stability criterion in Section IV, C, in conjunction with the maximum overlap principle.

Second, if one assumes that in the formation of a dative bond one electron is wholly transferred from the donor *D* to the acceptor *A*, the Sidgwick rule leads to assigning a very high effective charge to the central atom of complexes. For instance, in the case of ferrocyanide anion, one should have

$$[\mathrm{Fe(CN)_6}]^{4-} : Z(\mathrm{Fe}^{2+}) + 6 \text{ negative charges} = 30.$$

In other words, the iron atom should pass from the state of a dipositive ion Fe^{2+} to the state of a quadrinegative ion Fe^{4-}. To avoid writing such improbable charge transfers, Pauling (1948) proposed the following amendment to the Sidgwick rule: In stable complexes, the positive or negative net charge on each atom can hardly go beyond one electron; this recipe, called the "*electroneutrality principle*," means that any electron pair involved in a dative bond remains in the neighborhood of the donor atom almost entirely. As a matter of fact, the electric charge of an atom in a molecule is not a directly observable molecular property, but there is some experimental evidence of the electroneutrality principle: For instance, the isomeric shift of the iron atom in the Mössbauer spectra of ferro- and ferricyanides suggests that the iron atom is nearly neutral in all these compounds (Danon, 1963).

Theoretically, the effective atomic charges should be the result of an appropriate analysis of the total electronic wavefunction Ψ of the molecule under consideration (or more exactly of a partitioning process on the first-order matrix $\rho(1; 1')$ describing the electric cloud).

Nevertheless, as shown by Jørgensen (1963), it is possible to define the ionicity of an atom in a molecule and to estimate its magnitude in the frame of the classical "*electrochemical*" theories: If one takes for granted that the energy change required by the formation of a bond involving a charge transfer is derived from a sort of chemical potential, the value of which depends on the ionization potential of the donor atom and the electroaffinity of the acceptor atom, it appears that the bonded atoms try to adjust their charges in such a way that their electronegativities become more or less equal. Clearly, such a calculation method for effective charges leads to the electroneutrality principle, but lends itself to improvements moderating the effect of this rule: Correction terms can be incorporated from the beginning in the theory in order to take account of the additional stabilization energy due to the Coulomb interaction of the net atomic charges through the whole molecule, creating something known as an "*intramolecular Madelung potential.*"

B. The Electroneutrality Principle as a Result of the Equalization of Electronegativities

Let us consider an atom M surrounded by k ligands X, each of them carrying a net negative charge $-\delta$. In the most general case, the molecule MX_k may be neutral or charged and the ligands X identical or different atom groups. For reasons of simplicity, all the ligands will be assumed to be identical single atoms. In the case of a transition element M, the ligand X could be a halogen, as in the transition metal chlorides like $TiCl_4$ or $TiCl_6^{2-}$, or a chalcogen as in oxyanions MnO_4^- or CrO_4^{2-}.

Let us suppose that in a first stage we have to deal with a neutral compound MX_k. Then, the central atom carries a positive charge $+k\delta$ corresponding to a partial ionization of M (this quantity should be $-m + k\delta$ in the case of an oxyanion with a formal charge equal to $-m$).

In principle, atomic spectroscopy tables contain whatever is required to evaluate the extraction energy of a whole electron from a given shell of atom M (i.e., the various ionization potentials of M). A glance at the experimental values shows that this energy is not only dependent on the shell (nl), but also on the number q of electrons in it $(0 \leqslant q \leqslant 41 + 2)$, the occupation number of the other shells being assumed to be unchanged. For instance, in the case of titanium, the average ionization potential of a d electron, that is to say the energy difference between the baricenters of the initial and final configurations as defined by Slater (1960), is

$$I(\mathrm{Ti}, d^4) = E_{\mathrm{Av}}(\mathrm{Ti}^+, d^3) - E_{\mathrm{Av}}(\mathrm{Ti}, d^4) = 22{,}000\ \mathrm{cm}^{-1},$$

$$I(\mathrm{Ti}^+, d^3) = E_{\mathrm{Av}}(\mathrm{Ti}^{2+}, d^2) - E_{\mathrm{Av}}(\mathrm{Ti}, d^3) = 105{,}000\ \mathrm{cm}^{-1}.$$

Other configurations yield rather different values: 46,000 cm^{-1} and 140,000 cm^{-1} for the extraction of a d electron from $\text{Ti}(3d^3 4s)$ and $\text{Ti}^+ (3d^2 4s)$.

To estimate the energy of the molecular process producing a charge $+k\delta$ on atom M, it is then necessary to have a relationship between ionization energies and atomic charges. Many people have derived formulas of that sort, using linear or parabolic interpolation laws (see Jørgensen, 1962; Hinze *et al.*, 1963; Basch *et al.*, 1966a). The concept of "*differential ionization energy*" introduced by Jørgensen is perhaps the most significant of them: It is like a chemical potential acting in a molecule when a bond with charge transfer is formed (see Jørgensen, 1962, 1963, 1969, 1971a).

Given a sequence of ionization processes involving a shell (nl) occupied by q electrons without altering the occupation numbers of the other shells, the differential ionization energy is, by definition, an analytical function $I_{nl}(z)$ connected to the successive ionization potentials I_p of atom M by the expression

$$I_p = \int_{p-1}^{p} I_{nl}(z)\, dz \qquad (0 \leqslant p \leqslant q \leqslant 41 + 2)$$

where p is an integer taking the value 0 for the process $M^- \rightarrow M$ (I_0 corresponds to the electron affinity of M), the values 1,2, ... for the processes $M \rightarrow M^+$ (first ionization potential of M), $M^+ \rightarrow M^{2+}$, ... If as usual in atomic spectroscopy the variation of ionization potentials with respect to p is described by a parabolic law

$$I_p = b_0 + b_1 p + b_2 p^2$$

then the differential ionization is a quadratic expression in z

$$I_{nl} = a_0 + a_1 z + a_2 z^2,$$

the coefficients of which are obtained by integrating the latter expression between the limits p and $p - 1$ and identifying the results for $p = 0, 1, 2, \ldots$ to $I_0, I_1, I_2, \ldots$. Hence

$$a_0 = b_0 + \tfrac{1}{2} b_1 + \tfrac{1}{6} b_2, \qquad a_1 = b_1 + b_2, \qquad a_2 = b_2 .$$

Since ionization potentials are the energy change for two states of M differing in their total number of electrons, the integral of I_{nl} over z can be considered as a power series giving the expansion of the total energy of M in terms of its net charge z. Consequently, the integrand itself $I_{nl}(z)$ can be identified with the partial derivative $\partial E/\partial z$ of the energy, that is to say to an orbital electronegativity varying with the charge (Iczkowski and

Margrave, 1961; Hinze *et al.*, 1963). Of course, the analyticity of the function $I_{nl}(z)$ is simply an *ad hoc* hypothesis which has never been proved (Ferreira, 1967).

The coefficients a_0, a_1, a_2 are tabulated for the most typical configuration of each atom up to the sixth period (Jørgensen, 1969). They are not very different for a given subshell in closely related configurations, for instance the d electrons in $(3d^q)$, $(3d^{q-1}4s)$, $(3d^{q-1}4p)$ configurations of transition metals. As a general rule, the parameter a_2 is small with respect to the parameter a_1, which leads to a very simple relation between *differential ionization energies and electronegativities.* If the series expansions of ionization potentials are limited to the a_0 and a_1 terms, the ionization potential I_p becomes identical to the expression of the differential ionization energy $I_{nl}(z)$ for the value $z = p - \frac{1}{2}$:

$$I_p = I_{nl}(p - \tfrac{1}{2}) \qquad \text{if } a_2 = 0.$$

Therefore, the Mulliken generalized definition of electronegativity, namely the mean value of the energy change for two successive ionization processes I_p and I_{p+1} involving the electrons of some atomic shell (nl), gives:

$$\frac{I_p + I_{p+1}}{2} = \frac{I_{nl}(p - \frac{1}{2}) + I_{nl}(p + \frac{1}{2})}{2} = [I_{nl}(z)]_{z=p}.$$

At this approximation level, the electron affinity $A = I_0$ and the first ionization potential I_1 of the neutral atom are $I_{nl}(-\frac{1}{2})$ and $I_{nl}(+\frac{1}{2})$, respectively, giving a Mulliken electronegativity equal to $I_{nl}(0)$.

Let us now consider a closed system of several components, each of them being present in amount n_i. The chemical equilibrium condition states that

$$\sum_a \mu_a \, dn_a = 0$$

where μ_a is the chemical potential of the ath component, that is to say the partial derivative $(\partial E/\partial n_a)_{S,V,n_b}$ to be identified later with the derivative of the energy of each atom with respect to its net charge. Then, the "*chemical potential*" associated with an atom A whose electric charge is $\pm\delta$ should be written

$$\mu_A = (\partial E_A/\partial n_A)_{n_A = \pm\delta}.$$

To obtain an explicit expression for μ_A, it can be noted that the energy E_A of an atom with a whole number of electrons is given in terms of successive ionization potentials by the formula

$$E_A = -\sum_{i=-m,0,m}^{q} I_i = -\int_{-m,0,+m}^{q} I_{nl}(n_A)\, dn_A$$

where the quantities $q - m$, q, or $q + m$ are the number of electrons on the nl shell, according as A has the charge $+m$, 0 or $-m$. Assuming that this relation is still valid for fractional charges $\pm\delta$, the energy $E_A(\pm\delta)$ taken by the partial ion $A^{\pm\delta}$ is connected with that of the neutral atom $E_A(nl^q)$ by the equation

$$E_A(+\delta) = -\int_{\delta}^{q} I_{nl}(n_A)\, dn_A = E_A(nl^q) + \int_0^{+\delta} I_{nl}(n_A)\, dn_A,$$

$$E_A(-\delta) = -\int_{-\delta}^{q} I_{nl}(n_A)\, dn_A = E_A(nl^q) + \int_0^{-\delta} I_{nl}(n_A)\, dn_A,$$

from which the following expression is derived for the chemical potential of A:[1]

$$\mu_A = [I_{nl}(n_A)]_{\pm\delta}.$$

In the preceding expressions for $E_A(\pm\delta)$, the last integral is clearly the energy required to create a charge $\pm\delta$ on the nl shell of atom A, a result useful to compute the stabilization energy due to intramolecular charge transfers.

Let us suppose that the formation of the bonds in the molecule MX_k takes its origin in the charge transfer equal to $-\delta$ on each ligand X_a and $+k\delta$ on the central atom. The equilibrium condition for the chemical potentials of atoms M and X, namely

$$\mu_M\, dn_M + \sum_a \mu_{X_a}\, dn_{X_a} = 0$$

can be written

$$I_{nl}^M(+k\delta)\, dn_M + \sum_a I_{nl}^X(-\delta)\, dn_{X_a} = 0$$

where the variation of the charge dn_M on atom A is exactly equilibrated by the variation of the charges dn_{Xa} $(= dn_X)$ of atoms X:

$$dn_M + k\, dn_X = 0.$$

Consequently, the equilibrium condition for the whole molecule reduces to an equilibrium condition per bond

$$I_{nl}^M(+k\delta) - I_{nl}^X(-\delta) = 0.$$

---

[1] Note that the upper limit of the definite integral to be differentiated is a function of the variable.

In each bond M–X, the chemical potentials evaluated from the differential ionization energies of atoms M and X balance each other, so that the MX_k molecule has its energy minimum for a charge transfer δ satisfying the equilibrium condition. As a matter of fact, if one replaces I_{nl}^M and I_{nl}^X by their three-term expansion, one gets a quadratic equation in δ, the positive root of which gives the magnitude of the charge transfer δ. In the case of anions $[MX_k]^{-m}$, the same equation has to be solved assuming that the charge to be introduced into the expression of the differential ionization energy I_{nl}^M is $n_M = -m + k\delta$ instead of $+k\delta$.

The net charges given by this electrostatic model were calculated for a number of transition metal complexes (Jørgensen *et al.*, 1967; Jørgensen, 1969, 1971a). They do not exceed one electron on the central atom, except in oxyanions. For instance, one has

$$\begin{array}{ll} \mathrm{Ti}^{+0.88}\mathrm{Cl}^{-0.22} & \text{in } \mathrm{TiCl_4}, \\ \mathrm{Ti}^{+0.65}\mathrm{Cl}^{-0.44} & \text{in } \mathrm{TiCl_6^{2-}}, \\ \mathrm{Mn}^{+1.64}\mathrm{O}^{-0.66} & \text{in } \mathrm{MnO_4^{-}}, \\ \mathrm{Cr}^{+2.15}\mathrm{O}^{-1.04} & \text{in } \mathrm{CrO_4^{2-}}. \end{array}$$

In such calculations, small charge transfers are a general rule for neutral molecules as well as complex anions. From a certain point of view, they can be considered as a numerical materialization of the electroneutrality principle; from another point of view, they cannot be accepted as being really conclusive because the same low ionicity is also predicted for typical ionic compounds like NaCl (Jørgensen *et al.*, 1967). Nevertheless, the principle of electronegativity equalization or broadly speaking the hypothesis of a relationship between electronegativities and atomic net charges in molecules has often been used in semiquantitative descriptions of molecular structures (Haissinsky, 1946; Daudel and Daudel, 1946; Sanderson, 1952, 1960; Ferreira, 1967). For instance, the magnitude of charge transfers in organic or inorganic coordination compounds is fairly well predicted by a simple "*turning iteration method*" where the charges are connected to the effective electronegativities of atoms in molecules through the Pauling exponential formula giving the ionic character of a bond (Gallais *et al.*, 1965, 1966). The molecular orbital theory provides a quantum-mechanical background for semiempirical methods of that sort, to which we will come back in Section III, B. (For detailed discussions, see Klopman, 1964; Klopman and O'Leary, 1970; Hinze, 1968; Baird *et al.*, 1968; Van Hooydonk and Eeckhauf, 1970.)

C. The Effect of the Intramolecular Madelung Potential

A treatment related to the electroneutrality principle alone is deficient because it completely neglects the extra stabilization due to the interaction between the positive and negative charges created on the various atoms by the formation of bonds. On the analogy with the classical theory of ionic crystals, Jørgensen explained this additional attraction energy by the effect of the Madelung potential resulting from the intramolecular charge transfer and showed how the preceding development could be modified in order to take the *intramolecular Madelung energy* into account (Jørgensen *et al.*, 1967).

The electrostatic interaction of a system of point charges located on the atoms of a symmetrical molecule MX_k is (in atomic units)

$$E_{\text{elec}} = -\sum_a \frac{k\delta^2}{r_{M-X_a}} + \sum_{a<b}\sum \frac{\delta^2}{r_{X_a-X_b}}$$

where the bond length r_{M-X_a} has the same value, say r_{MX}, for the k ligands. This expression can be reduced to a Madelung form by writing

$$E_{\text{elec}} = -\frac{1}{2}\sum_a \frac{k\delta^2}{r_{M-X_a}} - \frac{1}{2}\sum_a \left(\frac{k\delta^2}{r_{M-X_a}} - \sum_{b\neq a} \frac{\delta^2}{r_{X_a-X_b}}\right)$$

and summing over the index a:

$$E_{\text{elec}} = -\frac{k}{2}\left[\frac{k\delta^2}{r_{MX}} - \left(\frac{k\delta^2}{r_{MX}} - \sum_{b\neq a} \frac{\delta^2}{r_{X_a-X_b}}\right)\right]$$

$$= -\frac{k}{2}(\alpha_M + \alpha_X)\frac{\delta^2}{r_{MX}}$$

with, by definition,

$$\alpha_M = k, \qquad \alpha_X = k - \sum_{b\neq a} \frac{r_{MX}}{r_{X_a-X_b}} < \alpha_M.$$

The Madelung coefficients α_M and α_X are completely determined by the conformation of the molecule. The numerical values of the ratio $r_{MX}/r_{X_a-X_b}$ for each distinct pair of ligands are given by simple geometrical considerations and tabulated up to MX_{12} for usual polyhedra (Jørgensen, 1969, 1971a). Therefore, a symmetrical MX_k molecule can be characterized by a Madelung constant

$$\alpha = \tfrac{1}{2}k(\alpha_M + \alpha_X) < k^2$$

and an electrostatic interaction energy of the form

$$E_{\text{elec}} = -\alpha(\delta^2/r_{MX}).$$

Let us suppose for the sake of convenience that the intramolecular Madelung energy just defined forms a part of the total energy of a central atom M having a net charge equal to $+k\delta$, and is written as

$$E_{\text{elec}}^{M} = -\frac{\alpha}{k^2}\frac{(k\delta)^2}{r_{MX}}$$

Then, the total energy of M becomes

$$E_M(+k\delta) = E_M(nl^q) + \int_0^{+k\delta} I_{nl}(n_M)\,dn_M - \left(\frac{\alpha}{k^2}\right)\left(\frac{n_M^2}{r_{MX}}\right)_{n_M = +k\delta}$$

It follows that the effective chemical potential of M is

$$\mu_M = \left(\frac{\partial E_M}{\partial n_M}\right)_{n_M = +k\delta} = I_{nl}^M(+k\delta) - \frac{2\alpha}{k}\cdot\frac{\delta}{r_{MX}},$$

and the equilibrium condition of the MX_k molecule including Madelung terms becomes

$$\left[I_{nl}^M(+k\delta) - \frac{2\alpha}{k}\frac{\delta}{r_{MX}}\right] dn_M + I_{nl}^X(-\delta)k\,dn_X = 0$$

with $dn_M = -k\,dn_X$. Consequently, at the energy minimum the charge transfer δ per bond will be given by the equation

$$I_{nl}^M(+k\delta) - I_{nl}^X(-\delta) - \frac{2\alpha}{k}\cdot\frac{\delta}{r_{MX}} = 0$$

where a Madelung term (the last one) will compensate the condition of electronegativity equalization $I_{nl}^M = I_{nl}^X$. As briefly indicated in Section II, B, the same charge calculation works in the case of a complex anion $[MX_k]^{-m}$ except for the fact that the central atom is not neutral from the beginning. Actually, the electron transfer from atom M produces an $(-m + k\delta)$ charge on M and therefore a Madelung energy increased by the quantity $+m/r_{MX}$. The complete equilibrium equation takes the form

$$I_{nl}^M(-m + k\delta) - I_{nl}^X(-\delta) + \frac{1}{r_{MX}}\left(m - \frac{2\alpha}{k}\delta\right) = 0.$$

The Madelung potential e/r_{MX} to be used for a given molecule is equal to 14.4 eV/Å. For the titanium compound previously mentioned, the following results are obtained:

$Ti^{+2.64}Cl^{-0.66}$ in $TiCl_4$ with $r_{TiCl} = 2.18$ Å,

$Ti^{+2.77}Cl^{-0.96}$ in $TiCl_6$ with $r_{TiCl} = 2.35$ Å.

As a general rule, the charge distributions found for transition metal complexes correspond to an intramolecular charge transfer much larger than the Pauling electroneutrality apparently predicts. In fact, they seem to be consistent with the gross atomic charges calculated by the molecular orbital method (see Section III,B), but their physical meaning is not exactly the same. First of all, the electrostatic model does not include any covalent-bonding effect, so that the stabilization energies computed by integrating the chemical potentials I_{nl}^{M} and I_{nl}^{X} over the charge distributions are too low (for instance, in $TiCl_4$, two thirds of the heat of formation is obtained). Second, the chemical bonding is related to a transfer of electrons in pure atomic orbitals $nl \in M$ and $n'l' \in X$ without any consideration of possible hybridization effects, which leads to disappointing results in the case of typical organic compounds (especially hydrides). To be sure, the preceding calculation method could be refined in such a way as to include some covalence and hybridization effects in the picture of the bond formation, but the main features of the electronic structure initially predicted are not dramatically altered by these improvements (Jørgensen, 1969, 1971a).

The main interest of the electrostatic model lies not in its numerical application to special molecules, but rather in the interpretation either of experimental data or of much more complicated quantum-mechanical treatments. As an illustrative example, let us consider magnesium difluoride MgF_2 in the usual linear form and in the bent form in which the molecule trapped in a krypton matrix at low temperature shows an apex angle $\theta = 158°$ (Calder *et al.*, 1969). The Madelung coefficient of the F atoms in the bent form is $\alpha_X = 2 - (2 \sin \theta/2)^{-1}$, that is to say 1.4906 instead of 1.5 in the linear form.

In the present case, the electrostatic model does not work very well because the equilibrium equation including the Madelung correction has a positive root giving a net charge $+k\delta$ outside of the domain of definition of the z variable for $Mg(3s^2)$. However, it is always possible to compute the difference of the electrostatic interaction energies in both forms, provided that the charge distribution of MgF_2 is known by another way. According to an *ab initio* molecular orbital calculation, the charge transfer δ is slightly larger in the linear form ($0.685e$) than in the bent form ($0.682e$) (Astier, 1970a). Taking $r_{Mg\text{-}F} = 1.77$ Å and $\delta = \frac{3}{4}$, one has:

$$\Delta E_{\text{elec}} = (3.5 - 3.4906)(\delta^2/r_{Mg\text{-}F}) = 1.0 \text{ kcal/mole}$$

in favor of the linear conformation. This value is in good agreement with the direct quantum-mechanical calculation ($\Delta E = 1.5$ kcal/mole), if the change of the correlation energy from one form to another is taken into account (Astier *et al.*, 1972). Clearly the electrostatic factors play an outstanding role in the electronic structure of MgF_2.[2]

III. Quantum-Mechanical Treatments of Coordination Compounds

A. The Application of the Molecular Orbital Method to Inorganic Chemistry

The classical theories of inorganic chemistry, as the electroneutrality principle and its refinements described above or even the crystal field theory discussed later in connection with the 10 *Dq* problem (Section IV,B), are mainly electrostatic models reducing a complex molecular system to a collection of free atoms or ions perturbed by external point charges. All of them completely neglect the fact that the ligands form true chemical bonds with the central atom, which should be treated by quantum mechanics. On the other hand, the molecular orbital method (LCAO–MO method) allows the study of the electronic structure of inorganic compounds with much more detail, since its formalism is able—at least in principle—to account for charge transfers due not only to electronegativity differences, but also to the nature of the atomic orbitals involved in chemical bonds or to some special conjugation effects (for instance, back-donation) between a central atom and its ligands.

The application of the LCAO–MO method to transition metal complexes has no important peculiarity, except that symmetry properties are systematically exploited in order to distribute the atomic orbitals of the ligands as well as those of the central atom among the symmetry species corresponding to the point group of the molecule, as it is intuitively done in the qualitative approach by the crystal field theory. The MO's are linear combinations of the form

$$\varphi_i(v) = \sum_{p=1}^{m} c_{pi}\, \chi_p(v)$$

where the $\chi_p(v)$'s are one-electron functions located on the various atoms of the molecule. In usual semiempirical treatments, the preceding expansion is limited to the atomic orbitals (AO) belonging to the valence shell of each atom, that is to say for a central atom of the iron series the $3d$, $4s$, and $4p$ AO's and for a ligand containing first-row atoms or hydrogens the

[2] This remark is due to Professor Jørgensen.

2*s* and 2*p* or 1*s* AO's. More complete sets of basis functions including inner-shell orbitals and involving other types of atomic wavefunctions (double-zeta functions, contracted Gaussian functions) have been used in recent calculations on metal complexes.

The c_{pi} coefficients of the *i*th linear combination φ_i are given by the usual set of homogeneous equations

$$\sum_{q=1}^{m} c_{qi}(H_{pq} - S_{pq} e_i) = 0 \qquad (p = 1, 2, \ldots, m),$$

the *i*th eigenvalue of which yields the energy of the electron allotted to φ_i. The matrix elements S_{pq} and H_{pq} are electronic integrals varying with each pair of AO's χ_p and χ_q. The overlap integrals.

$$S_{pq} = \int \chi_p{}^*(v)\chi_q(v)\, d\tau_v$$

can be evaluated in close form from the algebraic expression of the χ functions. The integrals

$$H_{pq} = \int \chi_p{}^*(v)H_{\text{eff}}(v)\chi_q(v)\, d\tau_v$$

are the matrix elements of an *effective Hamiltonian* for the motion of electron v in the average field of the other electrons. This one-electron Hamiltonian H_{eff} includes the kinetic and potential energy operators of v, the latter being the resultant of the attraction terms due to the nuclear framework and the average repulsion terms due to other electrons. In principle, the self-consistent field (SCF) theory leads to a well-defined expression of H_{eff} for closed-shell and open-shell systems (Roothaan, 1951, 1960), which could be used to evaluate the H_{pq} integrals numerically. However, for a long time it was practically impossible to apply the full SCF method to molecules containing atoms with orbitals of high quantum numbers, as is the case for transition-metal complexes. Now such calculations have been performed up to the iron series, but not for heavier atoms, and very often one must be satisfied with semiempirical formulas giving the numerical values of the H_{pq}'s in terms of the net atomic charges by possibly some sort of iterative process.

According to a recipe first suggested by Wolfsberg and Helmholz (1952) in a study of the ions MnO_4^- and CrO_4^{2-}, and in conformity with the ideas developed by Mulliken (1949), the diagonal elements H_{pp} are identified to the corresponding ionization energies of atom *P* (that is to say to the energy required for extracting one electron from the shell *nl* to which the orbital χ_p belongs), and the off-diagonal elements H_{pq} are estimated from the empirical values of the H_{pp}'s and the theoretical values of the overlap integrals S_{pq}. Several formulas have been recommended, namely the arithmetical rule (Wolfsberg and Helmholz, 1952),

$$H_{pq} = kS_{pq}(H_{pp} + H_{qq})/2,$$

the geometrical rule (Ballhausen and Gray, 1962),

$$H_{pq} = kS_{pq}(H_{pp} \cdot H_{qq})^{1/2},$$

and the reciprocal mean rule (Yeranos and Hasman, 1967),

$$H_{pq} = kS_{pq} \frac{H_{pp} \cdot H_{qq}}{(H_{pp} + H_{qq})/2}.$$

In all these formulas, the parameter k is a proportionality factor whose value may be fixed in relation to the nature of the orbitals χ_p and χ_q. The original values used by Wolfsberg and Helmholz were $k = 1.67$ for σ bonds and $k = 2.0$ for π bonds; they were chosen by fitting the k parameters on the experimental electronic transitions of oxyanions, the corresponding theoretical transitions being computed as simple orbital energy differences. Irrespective of the physical meaning of such calculations, one should not attach too much importance to the preceding values because the assignment of the various transitions in MnO_4^- and CrO_4^{2-} has been reconsidered since that time (see Oleari *et al.*, 1966; Dahl and Ballhausen, 1968; Astier, 1970b). As a matter of fact, the results obtained by means of the Wolfsberg–Helmholz formula or its substitutes do not depend too much on the constant k, qualitatively speaking, so that values going from 1 up to 4.5 have been successfully used: The standard value of k in the geometrical or reciprocal mean rule is $k = 2$ for all bonds. In the case of organic molecules involving σ and π bonds, a single value $k = 1.75$ also gave good results (Hoffmann, 1963).

This simplified MO method was very popular in the sixties, not only in theoretical inorganic chemistry, but also in organic chemistry, where it was called "*extended Hückel theory*" (EHT). However, a number of objections can be raised against its formalism:

(1) The diagonal elements H_{pp} do not always have a clear connection with the ionization potentials of atom P, because P should not be taken in its groundstate configuration, but rather in a valence state adapted—if possible—to the true electron distribution in the molecule under consideration, which is generally unknown at the beginning of the calculation. Iterative processes based on a relationship between the H_{pp}'s and the

orbital populations q_p of each atom are intended to correct this fundamental defect (see Section III, B).

(2) The approximations concerning the off-diagonal elements H_{pq} have very weak theoretical foundations. The arithmetical mean rule gives the same "*resonance integral*" H_{pq} for two orbitals strongly interacting because they correspond to two very close energy levels H_{pp} and H_{qq}, or for two weakly interacting orbitals obtained by adding and subtracting an arbitrarily large quantity to these energies, the half-sum $(H_{pp} + H_{qq})/2$ then being constant. The geometrical mean rule has not the same drawback, but overestimates the covalent character of the bond formed by the orbitals χ_p and χ_q. In addition, it contains an imaginary square root from the moment when the diagonal elements have opposite signs, forcing the product $H_{pp} \cdot H_{qq}$ to be taken in absolute value without any justification. In this respect, the reciprocal mean value seems to be better, predicting a bonding power in accordance with the best empirical measure of the covalence character (Allen, 1957).

It should also be noted that all the preceding formulas imply that the kinetic integral T_{pq} included in the matrix element H_{pq} is proportional to the overlap integral S_{pq}. Now kinetic terms are quadratic functions of the orbital exponents, so that a square should represent a better relationship between overlap and that part of the effective Hamiltonian (Cusachs, 1965). The best solution is probably to compute the matrix elements of the kinetic operator T from their definition in terms of the χ functions and to limit the approximations to the potential part V of the Hamiltonian using the arithmetic mean rule with $k = 1$ (Mulliken formula) for the multicentric integrals (Newton *et al.*, 1966).

(3) None of the above-mentioned formulas is invariant with respect to a linear transformation of the basis functions (for instance a hybridization of the atomic orbitals), nor a shift in the zero point of the energy scale, except for rather particular transformations (Löwdin, 1950; Berthier *et al.*, 1966). It is possible to construct semiempirical schemes which are invariant with respect to a rotation of the orbital coordinate axes by constraining theoretical parameters (integrals, matrix elements of H_{eff}, etc.) to fulfill *ad hoc* conditions (Pople *et al.*, 1965; Berthier *et al.*, 1966), but these assumptions may reduce the significance of the original method. In the case of the CNDO formalism, a strict rotational invariance is preserved by disregarding the angular part of the atomic orbitals in the calculation of integrals; in consequence, the method is unable to predict the shape of molecules stabilized by a strong π overlap [planar conformation of conjugated systems (Gropen and Seip, 1971), linear conformation of triatomic

molecules as MgF_2 (Astier *et al.*, 1972)]. Therefore, one could prefer a method able to give good predictions, even with a slight lack of invariance, to an unrealistic invariant model.

B. Relationship between the Hamiltonian Matrix Elements and the Net Atomic Charges

The difficulties reported in the previous section are of little importance if one accepts the use of semiempirical methods, not to perform absolute calculations on single molecules, but to study the trend of physical properties in related compounds. Actually, the general features of a series of molecules are hardly unchanged in any method if the parameters or the reference system for the orbitals are varied with discretion. On the other hand, the evaluation of the diagonal elements of the effective Hamiltonian H_{eff} makes much more trouble. Since the SCF Hamiltonian operator contains the atomic orbital populations implicitly, one can imagine various iterative MO methods where the elements H_{pp} are more or less complicated functions of the formal atomic charges. If the H_{pp}'s are still identified with atomic ionization energies, one is then forced to suppose that the data of atomic spectroscopy are continuous functions of the orbital populations of atoms, whether these populations are integers or fractional numbers, a problem which has been already discussed in Section II, B.

The calculation method developed by Basch *et al.* (1966a,b) can be taken as an example of semiempirical iterative scheme for transition metal complexes (for subsequent developments, see Fenske *et al.*, 1966; Fenske and Radtke, 1968). First, an average ionization potential for the shell $(nl)^q$ of an atom P bearing a formal electric charge δ is defined by the formula

$$I_{\text{VO}}(\delta) = I_{\text{O}}(\delta) + E_{\text{Av}}(P^{\delta}) - E_{\text{Av}}(P^{\delta+1})$$

where $I_{\text{O}}(\delta)$ is the ionization potential of P^{δ} in its spectroscopic ground-state and $E_{\text{Av}}(P^{\delta})$, $E_{\text{Av}}(P^{\delta+1})$ are the baricenters of the configurations involved in the ionization process with respect to their respective ground-states. These "*valence orbital*" ionization potentials (VOIP) have been tabulated for the first series of transition metals (Ballhausen and Gray, 1964; Basch *et al.*, 1966a). Next, the variation of I_{VO} for successive values of the formal charge δ on atom P is represented by a quadratic expression in δ:

$$I_{\text{VO}}(\delta) = a_0 + a_1\,\delta + a_2\,\delta^2$$

supposed to be also valid for fractional charges, and the function

$$H_{pp}(\delta) = -I_{\mathrm{VO}}(\delta) \qquad \text{for} \quad p \in (nl)$$

is used to construct an iterative scheme in terms of net atomic charges. Clearly, the VO ionization potentials $I_{\mathrm{VO}}(\delta)$ play the same role as the differential ionization energies $[I_{nl}(z + \frac{1}{2})]_{z=\delta}$ defined by Jørgensen (see Section II,B); both lead to the same charge picture, provided that the off-diagonal elements H_{pq} are calculated in a consistent way, as shown in Table I. Moreover, the total net charges given by such calculations for coordination compounds are *grosso modo* in agreement with the predictions of electrostatic models.

TABLE I

Diagonal Matrix Elements and Net Atomic Charges for $TiCl_4$ [a]

	$H_{pp} = I_{\mathrm{VO}}(\delta)^b$	$H_{pp} = I_{nl}(\delta + \frac{1}{2})^c$	$H_{pp} = I_{nl}(\delta + \frac{1}{2})$ + Madelung terms^c
$H_{3d\,3d}$	−90,800	−85,700	−98,700
$H_{4s\,4s}$	−79,500	−86,200	−74,000
$H_{4p\,4p}$	−45,300	−55,500	−38,400
q_{3d}^{Ti}	2.84	2.63	1.99
q_{4s}^{Ti}	0.19	0.40	0.29
q_{4p}^{Ti}	0.33	0.58	−0.28
δ(Ti)	+0.64	+0.34	+2.0

[a] Obtained with the geometrical mean rule ($K = 2.0$). Diagonal matrix elements in centimeters^{-1}.

[b] Basch *et al.* (1966b).

[c] Jørgensen *et al.* (1967).

In spite of their use, the parameters standing for the matrix elements H_{pp}, H_{pq} in semiempirical methods are not easily justified on theoretical grounds as long as the complete form of the effective Hamiltonian H_{eff} is not specified. For the present purpose, H_{eff} can be identified to the Fock one-electron Hamiltonian of the molecule

$$H_{\mathrm{eff}} \equiv F(\nu) = H^{\mathrm{core}}(\nu) + \sum_i [n_i J_i(\nu) - \tfrac{1}{2} n_i K_i(\nu)]$$

where n_i denotes the occupation number of each orbital $\varphi_i(n_i = 2, 1,$ or $0)$ and J_i, K_i are the Coulomb and exchange operators associated with φ_i in the SCF theory (Roothaan, 1951, 1960). Strictly speaking, such a simple expression is variationally correct only for closed-shell systems ($n_i = 2$ or 0), but it is not worthwhile to discuss the effect of the "*symmetry and equivalence restrictions*" (Nesbet, 1955) implied by its application to more general cases, if one is primarily concerned with possible improvements of semiempirical methods.

In 1965, a quasi-LCAO–SCF method based on the preceding Hamiltonian was developed for coordination complexes (Berthier *et al.*, 1965) using the Mulliken–Ruedenberg approximation in a systematic way to evaluate the bielectronic repulsion integrals contained in the matrix elements of the Coulomb and exchange operators J_i and K_i:

$$(pq;\cdot\cdot) \in J_i \simeq \tfrac{1}{2}S_{pq}[(pp;\cdot\cdot) + (qq;\cdot\cdot)],$$
$$(p\cdot;\cdot q) \in K_i \simeq \tfrac{1}{2}S_{pq}[(p\cdot;\cdot p) + (q\cdot;\cdot q)].$$

With these approximations, the diagonal element H_{pp} corresponding to an orbital χ_p of atom P becomes

$$\begin{aligned} H_{pp} = W_p &+ \sum_{r\in P} q_r^P(J_{pr} - \tfrac{1}{2}K_{pr}) \\ &+ \sum_{L\neq P}\Big[-(U_L{}^\circ\,;pp) - \sum_{l\in L}(n_l{}^L - q_l{}^L)(J_{pl} - \tfrac{1}{2}K_{pl})\Big] \end{aligned}$$

where $q_r{}^P$ and $q_l{}^L$ denote the "*gross atomic populations*" of orbitals χ_r and χ_l respectively centered on atoms P and L, that is to say expressions of the form

$$q_r{}^P = \sum_i \sum_s n_i c_{ir} c_{is} S_{rs} \quad \text{(sum over all basis functions } \chi_s)$$

(Mulliken, 1955), and $n_l{}^L$ is the number of electrons assigned to the orbital χ_l in the appropriate valence state configuration of atom L. The core parameters W_p and the penetration integrals $(U_L{}^\circ;\, pp)$ are straightforward extensions of the quantities defined by Goeppert-Mayer and Sklar (1938) for π-electron systems. Now, if one assumes that the formation of chemical bonds does not much alter the population of each orbital χ_l ($q_l{}^L \simeq n_l{}^L$) and, as usual in semiempirical treatments, one neglects all the penetration integrals, one finds

$$H_{pp} = W_p + \sum_{r\in p} q_r{}^P(J_{pr} - \tfrac{1}{2}K_{pr})$$

which is the basic formula for any iterative MO method. In its final form, the diagonal elements H_{pp} do not contain anything else than mono- and bielectronic integrals centered on atom P and can be evaluated, if possible, from atomic spectroscopy data for atom P and its ions. The integrals $J_{pr} \equiv (pp;\, rr)$ and $K_{pr} \equiv (pr;\, rp)$ have to be expressed in terms of Slater–Condon parameters F_k and G_k, and the core parameters W_p are computed by subtracting an appropriate combination of F_k and G_k parameters from experimental ionization potentials. The values of the parameter W_p and $(J_{pr} - \frac{1}{2}K_{pr})$ have been determined for the $3d$, $4s$, and $4p$ orbitals of the iron family (de Brouckère, 1970) and a complete list of parameters for the first and second transition-metal series has been recently published (Di Sipio *et al.*, 1971).

Clearly, the simplified expression found by putting $q_l^L \equiv n_l^L$ is valid only if the formation of the molecule does not require a too important charge transfer between the various atoms. In the case of strongly polar compounds, q_l^L is probably very different from n_l^L; consequently, it is not possible to drop the last term of H_{pp} and it becomes necessary to evaluate the bicentric interaction terms, at least through their long-range components J_{pl}^L. In fact, if one reduces all the bicentric Coulomb integrals to their asymptotic form

$$J_{pl}^{\mathrm{L}} \equiv (pp;\, ll) \simeq (1/R_{\mathrm{PL}}) \qquad \text{(in a.u.)},$$

R_{PL} being the distance between the atomic center of orbital χ_p and that of orbital χ_l, one introduces a correction term which is nothing but the Madelung potential created by the charge transfers from one atom to another, in conformity with the ideas of Jørgensen (de Brouckère, 1967). In practice, the $1/R$ approximation should be handled with caution, since the bicentric integral J_{pl} at small interatomic distances takes a finite value obviously related to those of the monocentric integrals J_{pp} and J_{ll}, and a formula of the Mataga–Nishimoto type

$$J_{pl} = \left(R_{\mathrm{PL}} + \frac{1}{\frac{1}{2}(J_{pp} + J_{ll})}\right)^{-1}$$

is more suitable for actual computations (de Brouckère, 1970).

The preceding method has been used in conjunction with the Wolfsberg–Helmholz approximation to calculate the electronic structure of various coordination compounds, including typical metallic complexes (Berthier *et al.*, 1965; de Brouckère, 1967) and large molecules (ferroporphyrine, vitamin B12 (Veillard and Pullman, 1965), bis (π-allyl) metal complexes (de Brouckère, 1970). A number of formalisms based on the same principles, but involving an evaluation of the off-diagonal elements at different approximation levels have been recently developed (see Nieuwpoort, 1965;

Manne, 1966, 1967; Canadine and Hillier, 1969; Hillier, 1969, 1970; Brown and Roby, 1970; Brown and Burton, 1970; Allen and Clack, 1970; Clack and Farrimond, 1971; Ciullio *et al.*, 1971; Tossell and Lipscomb, 1972).

C. Symmetry Considerations

Many usual coordination compounds have a rather high symmetry. Thus, it is useful to label the various orbital energy levels according to the irreducible representations of the point group G of the nuclear framework and to write the MO's themselves as linear combinations of symmetry orbitals. The symmetry orbitals, also called "*group orbitals*" $\xi_{g\alpha}$, are predetermined linear combinations of atomic orbitals:

$$\xi_{g\alpha} = \sum_r \lambda_{r\alpha} \chi_r$$

where the index g refers to a given set of geometrically equivalent orbitals χ_r (for instance the $2s$ orbitals of a set of equivalent ligands X_a) and the index α to one of the irreducible representations Γ_α of G spanned by the basis functions χ_r. Of course, symmetry-adapted basis functions are not an essential step when solving the LCAO–SCF equations, but the usual description of transition-metal complexes rests on symmetry considerations which are more easily understood in terms of MO's built from such functions: The various possibilities of mixing between ligand and metal orbitals and the distinction between σ and π bonds in octahedral and tetrahedral compounds are good examples.

The explicit form of the symmetry orbitals ξ can be found in the literature for many arrangements of ligands (see Kida *et al.*, 1957) but it should be borne in mind that the values of the coefficients λ are actually dependent on the direction of the reference axes x, y, z associated to each ligand. As a general rule, symmetry-adapted basis functions should be constructed by means of group-theory techniques using projection operators (see Serre, 1964). By definition, a symmetry orbital ξ belonging to an irreducible representation Γ_α of G is not altered if it is projected onto Γ_α itself. Therefore, if one projects out the various components χ of a set of geometrically equivalent orbitals by the standard formula

$$\mathrm{proj}_\alpha(\chi) = \frac{l_\alpha}{h} \sum_S R_S(\chi) U_\alpha^*(S)$$

—where h is the order of the point group G, l_α the dimension of the irreducible representation Γ_α, $R_S(\chi)$ the transformed function resulting from the effect of the symmetry operation S on an orbital χ and $U_\alpha(S)$ the character of S in Γ_α—one can write down the projection of any linear combination of the χ functions and choose the coefficients of the latter so that the whole combination remains invariant after a projection on Γ_α. The coefficients λ of the χ functions in a symmetry orbital ξ corresponding to a single representation are easily found by inspection. In the case of a two- or three-dimensional representation Γ_α (for instance the representations E and T of cubic groups), one has to construct two or three linearly independent combinations $\xi_{r\alpha}$ from the same set of equivalent orbitals χ_r. Unless these are obvious solutions (as for O_h and T_d symmetries), one has to replace the character U_α in the projection formula by appropriate matrix elements of the corresponding irreducible representation (Altmann, 1962).

The coordinate axes ordinarily used in qualitative descriptions, but not always in actual calculations (particularly by the promoters of general computer programs), are defined as follows: The orbitals of the central atom are referred to a right-handed coordinate system x, y, z and those of the ligands to so many left-handed systems, the z axis of which is pointing toward the central atom and the two others are in the perpendicular plane (see Ballhausen, 1962). With this assumption, the p_z orbitals of ligands can be considered as σ orbitals and the others as π ones and denoted by π, π', or π'' according as they are parallel to the x, y, or z axes of the general coordinate system attached to the central atom. Table II gives the symmetry species to which the MO's of a complex formula MX_6 or MX_4 belong for an AO basis set limited to the valence shells of the component atoms.

TABLE II

POSSIBLE TYPES OF MO'S IN MX_6 AND MX_4 COMPLEXES

Type[a]	O_h(27AO's)	T_d (21 AO's)	D_{4h} (21 AO's)
σ	$2 \times a_{1g}\,(s)$ $2 \times e_g\,(z^2, x^2 - y^2)$	$2 \times a_1\,(s)$	$3 \times a_{1g}\,(s, z^2)$ $2 \times b_{1g}\,(x^2 - y^2)$
$\sigma - \pi$	$3 \times t_{1u}\,(x, y, z)$	$4 \times t_2\,(x, y, z, xy, yz, xz)$	$3 \times e_u\,(x, y)$ $2 \times a_{2u}\,(z)$
π	$2 \times t_{2g}\,(xy, yz, xz)$	$2 \times e\,(z^2, x^2 - y^2)$	$2 \times b_{2g}\,(xy)$ $2 \times e_g\,(xz, yz)$
n (ligands)	$1 \times t_{1g}$ $1 \times t_{2u}$	$1 \times t_1$	$1 \times a_{2g}$ $1 \times b_{2u}$

[a] AO's of the central atom: 3d, 4s, 4p (in brackets). AO's of ligands: 2p.

In the case of highly symmetrical compounds, the main features of the electronic structure can be understood pretty well in terms of σ and π bondings. As an example, let us consider an octahedral complex MX_6, where the transition metal tries to approach the structure of krypton by adding 6 σ doublets taken from the ligands to its 18 core electrons and its q valence electrons; in addition, we suppose that each ligand contributes $2\pi_x$ and $2\pi_y$ electrons to the valence–electron system. For a noble gas structure ($q = 6$, as in complexes of Fe^{2+}), the system to be studied has 42 electrons and 27 basis AO's. First of all, the σ framework is formed by the 12 electrons assigned to the lowest energy levels $1a_{1g}$, $1e_g$, and $1t_{1u}$ (the latter being assumed to have a σ-predominant character). On the analogy with the directed valence theory, the mixing of two $3d$, one $4s$, and three $4p$ orbitals of M with the σ ligand orbitals in the symmetries a_{1g}, e_g, t_{1u} is often identified to a d^2sp^3 *hybridization* of M, but it should be added that the corresponding MO's have no directional character by themselves and that such an interpretation implies a relocalization of the standard LCAO–MO's (see Section IV,C). Above the preceding σ shells, π shells formed by the electrons of the $1t_{2g}$, $2t_{1u}$, $2t_{2g}$ levels and those of the $1t_{1g}$ and $1t_{2u}$ levels give some multiple-bond character to the bonding between the metal and the ligands, in spite of the fact that the higher level of t_{2g} symmetry is rather antibonding and the t_{1g} and t_{1u} levels are occupied by nonbonding electrons localized on ligands. Finally, the nature of the highest filled MO's and the lowest empty MO's plays a special role in the theoretical analysis of various molecular properties: The first ones occupied by the remaining q electrons ($q \leqslant 6$) belong to the π antibonding level $2t_{2g}$ previously mentioned; the second ones, completely empty in the ground state, correspond to a σ antibonding level $2e_g$ involving the two $3d_\sigma$ orbitals of M. The $2t_{2g}$ MO's are correlated with the π^* antibonding MO's which could be constructed for unsaturated ligands alone (like CO, CN^-, etc.), giving rise to the so-called "*back-donation*" effect, where a part of the q electrons initially located on the three $3d_\pi$ orbitals of M is transferred to the ligands, in the opposite direction to the σ charge transfer of the M–X bonds as requested by the electroneutrality principle. The empty $2e_g$ MO's are used to describe the first excited states of the molecule by jumping one electron from the highest occupied level t_{2g} to the lowest empty level e_g: The energy gap between these two levels is more or less understood as a splitting of the metal orbitals d_{xy}, d_{yz}, d_{zx} and d_{z^2}, $d_{x^2-y^2}$ due to the ligand field and related to the *parameter* 10 Dq of the crystal field theory (see Section IV,B).

Similar descriptions can be made, *mutatis mutandis*, for other types of transition-metal complexes; they work well as long as the molecular symmetry allows them to do so. This is the case of the main subgroups of the octahedron, save that the splitting of the *d* orbitals of the central atom is different: For instance, the threefold and twofold degenerate levels involved in the first electronic transitions are permuted in tetrahedral compounds and split into two in square-planar compounds. However, the number of σ-mixed MO's increases as the point group loses more and more elements, and the symmetry criteria become too poor to give a useful classification in terms of σ forward- and π backward-donation effects for the charge transfers and in terms of d_σ and d_π splittings for the electronic transitions. It is not excluded that a more palatable description may be obtained by applying a relocalization procedure to the canonical delocalized MO's.

D. *Ab Initio* Calculations of Metal Complexes

At the present time, there is a tendency to minimize the role of semiempirical methods in quantum chemistry and to prefer *ab initio* treatments as far as the nature of the molecule to be studied allows such an approach. This is indeed possible for simple coordination compounds containing one metal of the first transition series, as shown by the references quoted in Table III. Originally, it was expected that some doubtful aspects of approximate quantum-mechanical calculations could be avoided by *ab initio* methods where the molecular wavefunctions are derived from first principles, using basis functions only as input data. In practice, however, the form and the size of the atomic basis set in an LCAO expansion are more or less arbitrarily chosen, the number of AO's per atom being always limited in large molecules. The real meaning of the numerical results obtained by calculations of that sort may be disputed by people arguing that the basis functions are merely parameters as the α's and β's of the Hückel method. In any case, the ticklish question of the basis dependence of *ab initio* calculations has to be considered.

Certain quantities seem to be almost insensitive to the quality of the basis set, others are very much affected. The formal atomic charges belong to the first class: They are generally found to be in a reasonable agreement with earlier predictions based on classical or quantum models. For instance, in bis (π-allyl) complexes the nickel atom takes a positive charge $\delta_{\mathrm{Ni}} = +1.92$ (Veillard, 1969), as compared to the value $\delta_{\mathrm{Ni}} = +1.82$ obtained for a complex of the same symmetry (D_{2h}) by a simplified SCF method including Madelung-type correction terms (de Brouckère, 1970). Another striking example is the square-planar ion $[\mathrm{CuCl_4}]^{2-}$, an open-shell

TABLE III

Ab Initio SCF–MO CALCULATIONS FOR TRANSITION-METAL COMPLEXES

Molecules	Basis set[a]	Net Charges						Ref.[b]
$NH_3 \rightarrow ScH_3$	STO–MIN	Sc	+0.71	N	−0.56		—	1
		H	−0.28	H	+0.22		—	
$FTiH_3$	STO–MIN	Ti	+1.18	F	−0.28	H	−0.30	1
VO_4^{3-}	STO–3G/DZ	V	−0.07	O	−0.73		—	2
CrO_4^{2-}	STO–3G	Cr	+0.67	O	−0.68		—	3
—	STO–3G/DZ	Cr	+0.58	O	−0.64		—	2
CrO_8^{3-}	STO–2G/DZ	Cr	−0.72	O	−0.31	O	−0.26	4
—	GTO–MIN/DZ	Cr	+2.58	O	−0.70	O	−0.69	5
$Cr(CO)_6$	STO–3G/DZ	Cr	+0.70	C	+0.23	O	−0.35	6
MnO_4^-	STO–3G	Mn	+1.29	O	−0.57		—	3
—	STO–3G/DZ	Mn	+0.93	O	−0.48		—	2
—	STO–5G/7G	Mn	+1.33	O	−0.58		—	7
—	GTO–DZ	Mn	+1.11	O	−0.53		—	8
—	GTO–MZ	Mn	+0.99	O	−0.50		—	8
$Fe(C_5H_5)_2$	GTO–MIN/DZ	Fe	+1.23	C	−0.31	H	+0.19	9
$[NiF_4]^{2-}$	GTO–DZ	Ni	+1.18	F	−0.80		—	10
$[NiF_6]^{4-}$	GTO–MIN/DZ	Ni	+1.96	F	−0.99		—	11
—	GTO–MIN/DZ	Ni	+1.82	F	−0.97		—	12
$Ni(CO)_4$	STO–3G/4G/DZ	Ni	+0.47	C	+0.23	O	−0.35	6
$[Ni(CN)_4]^{2-}$	GTO–MIN/DZ	Ni	+0.46	C	−0.14	N	−0.47	13
—	STO–3G/4G/DZ	Ni	+0.81	C	−0.24	N	−0.46	14
$Ni(C_5H_5)NO$	STO–3G/4G/DZ	Ni	+0.82	C	−0.39	N	+0.07	14
				H	+0.23	O	−0.14	
$Ni(C_3H_5)_2$	GTO–MIN/DZ	Ni	+1.92	C from −0.64 to −0.34				15
				H from +0.22 to +0.08				
$[CuCl_4]^{2-}$	GTO–MIN/DZ	Cu	+1.05	Cl	−0.76		—	16

[a] STO–MIN: minimal set of Slater—type orbitals; STO–*n*G: minimal set of Slater-type orbitals fitted by n Gaussians; GTO–MIN: minimal set of contracted Gaussians; GTO–DZ: double-zeta set of contracted Gaussians; GTO–MZ: multiple-zeta set of contracted Gaussians.

[b] References to results of calculations published up to December, 1972:

(1) Stevenson and Lipscomb (1969).
(2) Connor *et al.* (1972).
(3) Hillier and Saunders (1970).
(4) Dacre and Elder (1972).
(5) Fischer *et al.* (1972).
(6) Hillier and Saunders (1971).
(7) Dacre and Elder (1971).
(8) Johansen (1972).
(9) Coutière *et al.* (1972).
(10) Basch *et al.* (1970).
(11) Gladney and Veillard (1969).
(12) Moskowitz *et al.* (1970).
(13) Demuynck *et al.* (1971).
(14) Hillier and Saunders (1972).
(15) Veillard (1969).
(16) Demuynck and Veillard (1970a).

system revealed by the study of crystals containing distorted octahedral complexes of Cu^{2+}: The SCF net charges given by the best *ab initio* calculation (Demuynck and Veillard, 1970a) coincide with the results of a valence–electron treatment (Trappeniers *et al.*, 1971) up to the second figure ($\delta_{Cu} = +1.05$ or $+1.07$), and the sequence of the orbital energies associated with the $3d$ electrons is comparable. To conclude definitely, one has to wait until *ab initio* results are available for directly observable quantities, as ESR and NQR data (Fermi contact term on Cu, anisotropy of the g tensor of the unpaired electron, nuclear quadrupole coupling constant of Cl), for which the valence–electron calculations are rather successful (Trappeniers *et al.*, 1972).

On the other hand, the decomposition of the various MO's in terms of metal and ligand components are subjected much more to basis effects. This is particularly disturbing in the case of the highest occupied and lowest empty MO's, since the usual description of metal complexes is based on the splitting of the d orbitals of metal among these orbitals. For instance, a preliminary *ab initio* calculation on the cluster $[NiF_6]^{4-}$ of $KNiF_3$ lattices gave an e_g orbital predominantly ligand in character for the MO occupied by two unpaired electrons (Hollister *et al.*, 1969), while an improved basis set yields a more pleasant picture with an open shell almost entirely metal (Moskowitz *et al.*, 1970).

IV. Present Status of Descriptive Models in Inorganic Chemistry

A. Crystal Field, Ligand Field, and Molecular Orbital Models

Nowadays, the traditional methods of theoretical inorganic chemistry (crystal-field and ligand-field theories, directed valence theory) are somewhat depreciated in consequence of the spread of general computer programs like POLYATOM or IBMOL.[3] Still, the interest of such approaches remains, because they allow one to "rationalize" a good deal of experimental and computational data which would be difficult to classify otherwise. In fact, they are models whose characteristics come from adjustments on a more general model, namely, the independent-electron atom, rather than on the physical reality. For mainly historical reasons, the crystal-field theory and the directed-valence theory are seen as opposite descriptions of the electronic structure of transition-metal complexes. It was

---

[3] Programs 92 and 199 of Quantum Chemistry Program Exchange, Indiana University (1973).

shown by Van Vleck that the total wavefunctions implied by the two descriptions are closely related (Van Vleck, 1935): The one corresponds to a completely reduced representation and the other to a reducible representation of the point group of the molecule in the space spanned by the valence orbitals of the central atom. The present report is only concerned with the connections between the various models, the details of which are to be found in standard textbooks.

The so-called "*crystal- and ligand-field*" (CLF) theory (Ballhausen, 1971) is a composite approach to the electronic spectrum of complexes based on the splitting of the d orbitals of the central atom M under the joint effect of electrostatic and covalence factors due to the binding with ligands X. In its simplest form, the CLF model can be visualized as a molecular orbital description using two-component MO's:

$$\begin{aligned}\varphi_+ &= N_+(\xi_{g\alpha} + \mu\chi_{M\alpha}),\\ \varphi_- &= N_-(\chi_{M\alpha} - \lambda\xi_{g\alpha}),\end{aligned}$$

for each irreducible representation Γ_α involved in the chemical bonds of a complex with symmetry G. The one-electron functions $\chi_{M\alpha}$ and $\xi_{g\alpha}$ are metal and ligand-group orbitals apt to mix with each other in the symmetry species Γ_α. The lower MO with a positive coefficient μ is a bonding orbital supposed to be essentially ligand in character ($\mu \ll 1$); its electrons belong to the valence shells of the molecule, but they are not explicitly considered in the CLF theory, being absorbed in atomic cores. The upper MO with a negative coefficient $-\lambda$ is an antibonding orbital usually associated with the d electrons of the metal M ($\lambda \ll 1$) whose splitting is studied. For instance, φ_+ and φ_- will represent either the orbitals $1e_g$ and $2e_g$ of octahedral compounds in a minimal set of valence orbitals (see Table II). The normalization factors N_+ and N_- are given by

$$N_+ = (1 + 2\mu S + \mu^2)^{-1/2}, \qquad N_- = (1 - 2\lambda S + \lambda^2)^{-1/2},$$

where S denotes a *group overlap integral*:

$$S = \int \xi^*_{g\alpha}(\nu)\chi_{M\alpha}(\nu)\, d\tau_\nu .$$

In the case of two-component MO's such as φ_+ and φ_-, the orthogonality of related bonding and antibonding MO's yields a strong condition between the values of λ and μ (Watson and Freeman, 1964)

$$\lambda = (\mu + S)/(1 + \mu S)$$

but the orthogonality constraints of the LCAO coefficients are much weaker in the case of extended CFL theories with several metal and/or ligand orbitals per MO.

The standard crystal-field theory is recovered by taking a purely ligand bonding orbital φ_+, that is to say by putting $\mu = 0$ and consequently $\lambda = S$ in the equations above. A first-order perturbation calculation based on a ligand-field potential V_c representing the electrostatic effect of ligands is made to evaluate the splitting of the d orbitals $\chi_{M\alpha}$ of the central atom. At this step of the theory, the antibonding MO

$$\varphi_- = (1 - S^2)^{-1/2}(\chi_{M\alpha} - S\xi_{g\alpha})$$

is not exactly a pure metal orbital, but rather a metal orbital orthogonalized to the ligand orbitals of the same species by the Schmidt procedure, and the ligand-field potential V_c should be identified with the difference between the SCF potential felt by the electron belonging to the molecular orbital φ_- and that of an electron in the atomic orbital $\chi_{M\alpha}$:

$$V_c(x, y, z) = V(x, y, z) - V_M(r),$$

the ligand electrons of the bonding orbital $\varphi_+ = \xi_{g\alpha}$ being incorporated in the core terms of the molecular potential $V(x, y, z)$ and the metal electrons being described by a central-field potential $V_M(r)$. As is well known, the overlap integrals S were originally neglected and the perturbation potential V_c was simulated by point charges δ_{X_a} located on ligand atoms, creating some sort of molecular *crystal field* which can be expanded in spherical harmonics around the central atom in order to estimate energy corrections. Thus, the first-order perturbation energy due to the crystal field

$$e_{CF}^{(1)} = \int \psi_-^*(v)V_c(v)\psi_-(v)\,d\tau_v = \int \chi_{M\alpha}^*(v)V_c(v)\chi_{M\alpha}(v)\,d\tau_v$$

is written as a sum of contributions such as

$$\int \chi_{M\alpha}^*(v)\,\frac{\delta_{X_a}}{\mathbf{r}(v) - \mathbf{r}_{X_a}}\,\chi_{M\alpha}(v)\,d\tau_v$$

and their computation in terms of Legendre polynomials leads to integrals of the form

$$I_n = \int_0^\infty R_{M\alpha}(r)\,\frac{r_<^n}{r_>^{n+1}}\,R_{M\alpha}(r)r^2\,dr$$

where $r_<$ and $r_>$ denote the lesser and the greater of the radius vectors $\mathbf{r}(\nu)$ and $\mathbf{r}_{X_a}$, and $R_{M\alpha}$ is the radial part of the atomic orbital $\chi_{M\alpha}$ (Bethe, 1929; Van Vleck, 1932, 1939; Schlapp and Penney, 1932; Ilse and Hartmann, 1951). The integrals I_n and their first and second derivatives with respect to the bond distance r_{X_a}, which are useful for studying the effect of dipolar charge distributions and Jahn–Teller distortions, have been tabulated for the values of n involved by d electrons, namely $n = 0, 2, 4$ (Ballhausen and Ancmon, 1958).

A more correct theory should make provision for nonzero λ coefficients. In principle, the values assigned to both coefficients would be immaterial for a pair of completely filled two-component MO's like ψ_+ and ψ_-, since a closed-shell Slater determinant is invariant with respect to any unitary transformation of the basis MO's (Watson and Freeman, 1964). Nevertheless, the magnitude of λ becomes important in the case of a transition-metal complex with unpaired electrons in the d shell, because it is connected with the hyperfine structure of its ESR spectrum through the spin-density parameters f_s and f_σ on ligand nuclei (see McGarvey, 1967). Therefore, it is worthwhile to consider a first-order CLF theory based on the assumptions

$$\mu = 0, \qquad \lambda = S \neq 0;$$

in other words to reintroduce the overlap integrals in the primitive crystal-field theory with intent to take ligand effects into account. Then, the perturbation energy due to the crystal and ligand field is

$$\begin{aligned} e_{\mathrm{CLF}}^{(1)} &= \int \psi_-^*(\nu)V_c(\nu)\psi_-(\nu)\,d\tau_\nu \\ &= (1 - S^2)^{-1}\left[\int \chi_{M\alpha}^*(\nu)V_c(\nu)\chi_{M\alpha}(\nu)\,d\tau_\nu - 2S\int \chi_{M\alpha}^*(\nu)V_c(\nu)\xi_{g\alpha}(\nu)\,d\tau_\nu \right. \\ &\quad \left. + S^2\int \xi_{g\alpha}^*(\nu)V_c(\nu)\xi_{g\alpha}(\nu)\,d\tau_\nu\right]. \end{aligned}$$

The preceding expression can be written as a series expansion with respect to the overlap integral S and tentatively limited to the second-order terms in S. Using the Mulliken–Ruedenberg approximation (Ruedenberg, 1951) to calculate the value of the function $\chi_{M\alpha}(\nu)$ in the neighborhood of the ligand nuclei,

$$\chi_{M\alpha}(\nu) \simeq S\xi_{g\alpha}(\nu) \qquad (\mathbf{r}_\nu \simeq \mathbf{r}_{X_a}),$$

one obtains

$$e^{(1)}_{\text{CLF}} \simeq (1 + S^2)\left[\int \chi^*_{M\alpha}(v)V_c(v)\chi_{M\alpha}(v)\, d\tau_v - S^2 \int \xi^*_{g\alpha}(v)V_c(v)\xi_{g\alpha}(v)\, d\tau_v\right]$$

$$\simeq e^{(1)}_{\text{CF}} + S^2(e^{(1)}_{\text{CF}} - e^{(1)}_{\text{LF}})$$

where the quantity

$$e^{(1)}_{\text{LF}} = \int \xi^*_{g\alpha}(v)V_c(v)\xi_{g\alpha}(v)\, d\tau_v$$

represents an intrinsic ligand-field contribution (Ballhausen, 1971). The second term is proportional to the integral S squared and gives an angular dependence of the d splitting of the central atom which is a function of the position of the ligands. Even if this fact is of slight consequence in a semi-empirical theory where $e^{(1)}_{\text{CF}}$ or $e^{(1)}_{\text{CLF}}$ are considered as adjustable parameters, it plays an important role in subsequent improvements of the ligand-field theory (Jørgensen *et al.*, 1963).

To develop a second-order CLF theory, it is necessary to specify the form of the diagonal and off-diagonal CLF Hamiltonian in a little more detail. Let us consider a one-electron effective Hamiltonian of the form

$$H_{\text{eff}}(v) = -\tfrac{1}{2}\nabla^2(v) + V_M(v) + V_X(v)$$

in which the potential acting on the electron v has been split into a metal contribution V_M and a ligand contribution V_X corresponding to the two components $\chi_{M\alpha}$ and $\xi_{g\alpha}$ of $\varphi_\pm$.

If the electron v would be submitted to the field of M or X alone, its energy would be

$$e^{(0)}_M = \int \chi^*_{M\alpha}(-\tfrac{1}{2}\nabla^2 + V_M)\chi_{M\alpha}\, d\tau,$$

$$e^{(0)}_X = \int \xi^*_{g\alpha}(-\tfrac{1}{2}\nabla^2 + V_X)\xi_{g\alpha}\, d\tau.$$

The energy levels e are obtained by solving the 2 × 2 secular equation

$$(H_{MM} - e)(H_{XX} - e) - (H_{MX} - eS)(H_{XM} - eS) = 0$$

where H_{MM}, H_{MX}, and H_{XX} are the matrix elements of the one-electron Hamiltonian H_{eff} with respect to the basis functions $\chi_{M\alpha}$ and $\xi_{g\alpha}$.

The partitioning technique (Löwdin, 1964) provides an elegant way of formulating the problem.[4] The various elements of the secular determinant

---

[4] This was suggested by Professor Löwdin (Sanibel Lectures).

W_{AB} can be considered as functions of the energy parameter e:

$$W_{AB}(e) = H_{AB} - eS_{AB}$$

leading to a secular equation of the form

$$W_{MM} - W_{MX} W_{XX}^{-1} W_{XM} \doteq 0$$

whose solutions can be written $e = f(e)$ by letting

$$f(e) = H_{MM} - (H_{MX} - eS)(H_{XX} - e)^{-1}(H_{XM} - eS).$$

The coefficients C_M and C_X of the basis functions verify the following equations:

$$W_{MM} C_M + W_{MX} C_X = 0, \qquad W_{XM} C_M + W_{XX} C_X = 0.$$

If one takes it for granted that the function φ_- previously defined is an eigenfunction of the problem, one has

$$C_M = (1 - S^2)^{-1/2}, \qquad C_X = -S(1 - S^2)^{-1/2},$$

and consequently

$$(H_{XM} - eS) - (H_{XX} - e)S = 0.$$

At this approximation level, the secular equation is reduced to the simple form

$$(H_{MM} - e) - (H_{MX} - eS)S = 0$$

whose solution is

$$e = (1 - S^2)^{-1}(H_{MM} - H_{MX} S).$$

The values of the matrix elements H_{MM} and H_{MX} are

$$H_{MM} = e_M^{(0)} + \int \chi_{M\alpha}^* V_X \chi_{M\alpha}\, d\tau,$$

$$H_{MX} = e_M^{(0)} S + \int \chi_{M\alpha}^* V_X\, \xi_{g\alpha}\, d\tau,$$

the latter expression supposing, at least theoretically, that the orbital χ_M is an eigenfunction of the operator $(-\frac{1}{2}\nabla^2 + V_M)$.[5] One gets

$$e = e_M^{(0)} + (1 - S^2)^{-1}\left(\int \chi_{M\alpha}^* V_X \chi_{M\alpha}\, d\tau - S \int \chi_{M\alpha}^* V_X \xi_{g\alpha}\, d\tau\right),$$

[5] If it is not the case, even *grosso modo*, it should be probably better to expand H_{MX} symmetrically with respect to V_M and V_X (Sender and Berthier, 1958; Dunn, 1970).

that is to say an expression entirely identical to the preceding perturbation formula with overlap provided that the coefficient $(1 - S^2)^{-1}$ and the last integral are expanded in a series in S^2.

A second-order approximation (of course valid only if $|H_{MX}| \ll |H_{MM} - H_{XX}|$) can be derived by introducing the quantity $e_M^{(0)}$ into the partitioning formulas instead of the energy e of the metal orbital in the field of the ligands. For the numerator of the second-order term one finds

$$H_{MX} - eS = \int \chi_{M\alpha}^* V_X \xi_{g\alpha} \, d\tau$$

and for the denominator one has

$$H_{XX} - e \simeq e_X^{(0)} - e_M^{(0)}$$

where

$$e_X^{(0)} = \int \xi_{g\alpha}^*(-\tfrac{1}{2}\nabla^2 + V_X)\xi_{g\alpha} \, d\tau.$$

Thus, a metal electron in a ligand field V_X takes an energy of the form

$$e = e_M^{(0)} + \int \chi_{M\alpha}^* V_X \chi_{M\alpha} \, d\tau + \left| \int \chi_\alpha^* V_X \xi_{g\alpha} \right|^2 \Big/ (e_M^{(0)} - e_X^{(0)}) + \cdots$$

where the first term $e_M^{(0)}$ is the orbital energy of $\chi_{M\alpha}$ in the isolated atom. The second term

$$e_M^{(1)} = \int \chi_{M\alpha}^* V_X \chi_{M\alpha} \, d\tau$$

coincides with the crystal-field correction $e_{\mathrm{CF}}^{(1)}$, the third

$$e_M^{(2)} = \left| \int \chi_{M\alpha}^* V_X \xi_{g\alpha} \, d\tau \right|^2 \Big/ e_M^{(0)} - e_X^{(0)}$$

includes a part of the CLF correction previously defined,[6] and the quantity $\Delta e_M = e_M^{(1)} + e_M^{(2)}$ gives the shift of the energy levels of M as a result of the ligand field.

The preceding second-order CLF theory, initially developed by Jørgensen *et al.* and called "*angular overlap model*" (AOM) yields energy shifts

---

[6] An expression equivalent to the Ballhausen correction term $e_{\mathrm{CLF}}^{(1)}$ is obtained by substituting the diagonal element H_{MM} for the energy e in the partitioning formula, but it includes a term proportional to the fourth power of the $F_{M,\lambda}$ coefficient.

having a dependence of the overlap type with respect to the angular part of the metal orbital $\chi_{M\alpha}$ and to the arrangement of the ligands X_a (Jørgensen *et al.*, 1963; Schaeffer and Jørgensen, 1965a,b). In fact, a group overlap integral S can be expressed as follows:

$$S = \Xi_\lambda(\chi_M, X_1, X_2, \ldots, X_N)S_{\mathrm{rad}}$$

where S_{rad} is a radial group integral computed from nodeless S functions (renormalized by the factor $(4\pi)^{-1/2}$) instead of the true metal and ligand orbitals. The subscript λ attached to the angular part Ξ refers to the different species of chemical bonding resulting from the overlap between metal and ligand orbitals, namely σ, π and δ for $\lambda = 0$, 1, and 2 in the case of d metal orbitals. To calculate the contributions of type π, δ, etc., a double index $\lambda\omega$ corresponding to the two-dimensional representations of the local symmetry group of the bond M–X_a may be necessary, and the overlap integral S is obtained by adding the various ω contributions if the group orbital $\xi_{g\alpha}$ includes components of both sorts (for instance in a symmetry orbital t_{2g}^{xy} of O_h, the orbitals noted π_x and π_y with respect to the reference axes of M).

Usually, the Ξ coefficients are written

$$\Xi_{\lambda\omega}(M, X_a) = \mathcal{N}_\lambda F_{M,\lambda\omega}(\theta_a, \varphi_a)$$

where (θ_a, φ_a) are the angular coordinates of the ligand X_a, and the proportionality factor $\mathcal{N}_\lambda$—equal to $(2l + 1)^{1/2}$ for σ bonding with an l orbital χ_M, $[2(2l + 1)l(l + 1)]^{1/2}$ for π bonding, $[\frac{1}{2}(2l + 1)(l - 1)l(l + 1)(l + 2)]^{1/2}$ for δ bonding ($\mathcal{N}_\lambda = 5^{1/2}$ for σ bonding or $2(15)^{1/2}$ for π and δ bondings with d orbitals)—is incorporated into the radial part of S, which becomes

$$S = F_{M,\lambda}(\theta_a, \varphi_a)\tilde{S}_{MX_a},$$

the sign over the overlap integral S_{MX_a} denoting the absence of angular parts in the atomic functions. The F coefficients have been tabulated in trigonometric form (Schaeffer and Jørgensen, 1965a,b; Schaeffer, 1970); their squares are simple fractional numbers, the values of which are often found by inspection of the Cartesian expression $A_l(x, y, z)/r^l$ of the angular part of χ_M.

It is possible to factorize the ligand-field integrals involving the potential V_X in the same way, provided that the three-center integrals arising from the expansion of V_X in terms of ligand-core contributions

$$V_X = \sum_a V_{X_a}(\nu)$$

are neglected in the second-order term (Schmidtke, 1971). Using the F coefficients, one gets

$$\int \chi_{M\alpha}^{*} \sum_{a} V_{X_a} \chi_{M\alpha}\, d\tau = \sum_{a} F_{M,\lambda}^{2}(\theta_a, \varphi_a)\, \tilde{V}_{MM},$$

$$\int \chi_{M\alpha}^{*} \sum_{a} V_{X_a} \xi_{g\alpha}\, d\tau \simeq \int \chi_{M\alpha}^{*} \sum_{a} V_{X_a} \chi_{X_a}\, d\tau,$$

$$\int \chi_{M\alpha}^{*} \sum_{a} V_{a} \chi_{X_a}\, d\tau = \sum_{a} F_{M,\lambda}(\theta_a, \varphi_a)\, \tilde{V}_{MX_a},$$

the tilde over the ligand-field integrals of both types having the same meaning as for overlap integrals. If in addition one takes $e_X^{(0)} = e_{X_a}^{(0)}$, the shift of the energy level $e_M^{(0)}$ can be written

$$\Delta e_M = \sum_{a} F_{M,\lambda}^{2}(\theta_a, \varphi_a) e_{\lambda, X_a},$$

the quantity e_{λ, X_a} defined by

$$e_{\lambda, X_a} = \tilde{V}_{MM} + [\tilde{V}_{MX_a}^{2}/(e_M^{(0)} - e_{X_a}^{(0)})]$$

being a parameter characteristic of the M–X bond, whatever the molecular environment is. If the atoms M and X have appropriate orbitals χ, there are different parameters $e_\sigma, e_\pi, e_\delta \ldots$ for each possible type of bonding between M and X.

The reduction of the second-order ligand-field theory to a single set of parameters e_λ playing the same role as first-order crystal-field splittings $e_M^{(1)}$ suggests that a correspondence should exist between the predictions of both models for a compound of given symmetry. Actually, a close relationship between the parameters e_λ of the angular model and the integrals I_n of the crystal-field theory can be established on symmetry grounds (Schaeffer and Jørgensen, 1965b; Kibler, 1970). For instance, in the case of orthoaxial octahedral complexes, one has

$$e_\sigma - e_\delta \propto \tfrac{4}{7} I_2 + \tfrac{5}{21} I_4,$$

$$e_\pi - e_\delta \propto \tfrac{3}{7} I_2 - \tfrac{5}{21} I_4.$$

Thus, the two models are practically equivalent as long as their respective parameters can be considered as freely adjustable quantities. However, the values of the crystal-field splittings $e_M^{(1)}$ which would be necessary to explain the absorption spectra of transition-metal complexes (see Section IV,B below) have almost nothing to do with the results of quantum-mechanical calculations based on the same model and are strongly dependent on the molecular environment. In this respect, the angular model apparently works better, for it involves parameters which are transferable by definition from one compound to another, irrespective of the symmetry, and take overlap and covalence effects into account. It should be also added that the preceding analysis based on the partitioning technique should be most likely applicable in the case of more general CLF models with more than one orbital per atom, since the elements W_{AB} can be matrices instead of single numbers.

B. The $10Dq$ Problem

The first object of ligand-field theories is to explain why many inorganic compounds absorb the light in the near ultraviolet or the visible region. As the q valence electrons of the central atom M in a transition-metal complex cannot generally fill up its *nd* shell ($q \leqslant 10$), the lowest absorption bands of the optical spectrum are ascribed to electronic transitions inside this valence shell, the *nd* level of M being split into several components by the ligand field. A classical example is the case of symmetrical compounds MX_6, where the degeneracy of the d level is lifted by the effect of an octahedral ligand field giving rise to a lower level t_{2g} and an upper level e_g. The magnitude of the energy gap $(e_g - t_{2g})$, the value of which is given in the standard crystal-field theory by the expression[7]

$$e_M^{(1)}(d_{z^2}, d_{x^2-y^2}) - e_M^{(1)}(d_{xy}, d_{yz}, d_{xz}) = 6Dq - (-4Dq) = 10Dq,$$

determines the filling of the two levels in the groundstate configuration. If the gap is small enough (weak-field limit), the t_{2g} level and then the e_g level are occupied by the q electrons in such a way that the resultant spin is maximum, as if t_{2g} and e_g would be degenerate. If the gap is large enough (strong-field limit), the t_{2g} level is first occupied by the greatest possible number of electrons, in conformity with the prescriptions of the Pauli principle and of the Hund rules, and afterward the e_g level. Characteristic absorption bands are due to the excitation of an electron from the t_{2g} level to the e_g level.

The expression of the $(e_g - t_{2g})$ gap in the angular model is immediately obtained by writing the d orbitals in the form

$$d_{z^2} \propto z^2 - \tfrac{1}{2}(x^2 + y^2), \qquad d_{x^2-y^2} \propto (3^{1/2}/2)(x^2 - y^2), \qquad d_{xy} \propto xy, \quad \text{etc.}$$

---

[7] For a definition of the so-called "*cubic-field strength parameter*" Dq, see Schläfer and Glieman (1969).

The axial overlap of the d_σ orbitals with group orbitals $\xi_{X\alpha}$ of σ type gives a shift equal to

$$\begin{aligned}\Delta e_{M,\sigma} &= [F_\sigma{}^2(z^+) + F_\sigma{}^2(z^-) + F_\sigma{}^2(x^+) + F_\sigma{}^2(x^-) + F_\sigma{}^2(y^+) + F_\sigma{}^2(y^-)]e_\sigma \\ &= (1 + 1 + \tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{4})e_\sigma = 3e_\sigma\end{aligned}$$

for the $3d_{z^2}$ orbital and

$$\begin{aligned}\Delta e_{M,\sigma} &= [F_\sigma{}^2(x^+) + F_\sigma{}^2(x^-) + F_\sigma{}^2(y^+) + F_\sigma{}^2(y^-)] \\ &= (\tfrac{3}{4} + \tfrac{3}{4} + \tfrac{3}{4} + \tfrac{3}{4})e_\sigma = 3e_\sigma\end{aligned}$$

for the $3d_{x^2-y^2}$ orbital, one ligand being located on the positive side $x^+y^+z^+$ and another on the negative side $x^-y^-z^-$ of each coordinate axis. The lateral overlap of the d_π orbitals with the group orbitals $\xi_{X\alpha}$ of π type gives

$$\begin{aligned}\Delta e_{M,\pi} &= [F_\pi{}^2(x^+) + F_\pi{}^2(x^-) + F_\pi{}^2(y^+) + F_\pi{}^2(y^-)]e_\pi \\ &= (1 + 1 + 1 + 1)e_\pi = 4e_\pi\end{aligned}$$

for the $3d_{xy}$ orbital, and similar contributions for the orbitals $3d_{yz}$, $3d_{xz}$. Then, in the angular model the energy gap $(e_g - t_{2g})$ takes the form (Schaeffer and Jørgensen, 1965b; Schmidtke, 1971)

$$10Dq = 3e_\sigma - 4e_\pi .$$

In fact, if the ligands have atomic orbitals of appropriate symmetry, there are additional contributions of δ type to the splitting of both levels ($+3e_\delta$ for e_g, $-2e_\delta$ for t_{2g}), but their net contribution to the energy gap, $+e_\delta$, can be incorporated into the σ and π parts by putting

$$\begin{aligned}&e'_{\lambda, X_\alpha} = e_{\lambda, X_\alpha} - e_{\delta, X_\alpha} \qquad (\lambda = \sigma \text{ and } \pi), \\ &10Dq = 3e_\sigma{}' - 4e_\pi{}' .\end{aligned}$$

Unlike the crystal-field theory, the angular model uses a bidimensional series of parameters to express the transition energies, but this peculiarity does not play any role for the moment there is only one experimental quantity to be interpreted.

As is well known, similar relations hold for tetrahedral complexes, except that the levels e and t are reversed and the energy gap is smaller (Jørgensen, 1971a):

$$(e - t_2)_{\text{tet}} = -\tfrac{3}{4}e_\sigma{}' + \tfrac{16}{9}e_\pi{}' = -\tfrac{4}{9}(e_g - t_{2g})_{\text{oct}}$$

(the bond distance M–X being assumed to be equal in tetrahedral and octahedral compounds). Extra splittings occur when the molecular symmetry lowers, for instance in square-planar complexes.

Many theoretical works have been devoted to the evaluation of the quantity $10Dq$, in connection with the study of the first excited states in transition-metal complexes. The question is complicated by the fact that, as a general rule, the transition energy Δ between two states of well-defined symmetry and multiplicity cannot be reduced to a simple energy gap, even in the frame of an independent-particle model using the same set of MO's in both states because the orbital- and/or spin-coupling conditions of the electron system vary, except in very special cases, when passing from the initial state to the final state; bielectronic interaction integrals have to be added to the energy difference of the orbitals involved in the excitation process. This difficulty also arises in CLF theories for the very reason that the q electrons of the valence shell of the central atom are treated as a whole, the valence electrons of the ligands only being hidden together with the inner shells in atomic cores.

A rather general expression for transition energies inside the valence shell $(nd)^q$ of an octahedral compound is

$$\Delta = I_{e_g} - I_{t_{2g}} + \sum_{i,j,k,l} \lambda_{ijkl} G_{ijkl}$$

where the quantity $(I_{e_g} - I_{t_{2g}})$ has to identified with $10Dq$ and the G's are bielectronic integrals with indices $ijkl$ corresponding to the three t_{2g} orbitals and the two e_g orbitals. The crystal-field theory in its simplest form neglecting overlap reduces the t_{2g} and e_g functions to pure nd orbitals of metal; thus, the G's become one-center bielectronic integrals which can be expressed in terms of Slater–Condon parameters $F^k(dd)$, or alternatively in terms of Racah parameters A, B, C.[8] The coefficients λ_{ijkl} are obtained in calculating each transition energy as a difference of total energies between two states with well-defined symmetry and multiplicity. To determine the necessary wavefunctions, one must combine the Slater determinants representing the same configuration in an appropriate way and mix them with other $(nd)^q$ configurations of the same symmetry if configuration interaction has to be taken into account. The matrix elements required by such a calculation have been tabulated, including spin-orbit corrections (Griffith, 1961). As a general rule, the Racah parameter A does not appear in the expression of Δ, but the parameters B and C remain, except in particular cases (for instance in the transition ${}^3T_{2g} \leftarrow {}^3A_{2g}$ of d^8 complexes, like $[\mathrm{NiF_6}]^{4-}$). The results have been collected in the form of diagrams (Tanabe and Sugano, 1954; Orgel, 1955) for d^q configurations:

---

[8] $A = F^0 - (49/411)F^4$, $B = (1/49)F^2 - (5/411)F^4$, $C = (35/441)F^4$.

The energies of the excited states are plotted in terms of the parameters $10Dq$ and B, the values of which can be fitted on experimental spectra. It is found that the best Racah parameters are slightly smaller in a complex ion than in a free one, and the reduction of B called "*nephelauxetic effect*" is interpreted as an expansion of the electron cloud of the metal toward the ligands (Schaeffer and Jørgensen, 1958), but at this stage of the theory, the effective value of the gap $10Dq$ is an empirical quantity whose origin is still obscure.

Before translating the parameter $10Dq$ in the language of the molecular orbital theory, it is necessary to make precise the meaning of the energy attributed to an electron on a given spin orbital. In an N-electron system described by an independent-particle model, the one-electron energy levels are usually defined by an expression of the form

$$e_i = I_i + \sum_j J_{ij} - \sum_j' K_{ij}.$$

The first term I_i includes the kinetic energy and the nuclear attraction energy of the electron assigned to the spin orbital ψ_i, the last quantity being possibly corrected by the repulsion energy coming from the electrons hidden in the atomic cores. The second and third terms collect a number of bielectronic integrals G with only two different indices, namely the Coulomb and exchange integrals J_{ij} and K_{ij}, which express the electronic interaction within the system in a certain way. Actually, in the classical SCF method for closed-shell systems, the sum $\sum_j$ is taken over all the *spin orbitals* ψ_j occupied in the ground state, whereas the sum $\sum_j'$ is restricted to those spin orbitals ψ_j having the same spin part as ψ_i. The same summation rules are used for a virtual spin orbital to be occupied in an excited state, and this plays an important role when expressing the transition energies in terms of orbital energies (Roothaan, 1951; Dahl and Ballhausen, 1968). Irrespective of the difficulties peculiar to the treatment of open-shell systems by the SCF–MO method (Berthier, 1964), the preceding definition leads to serious inconsistencies as concerns the formulation of the $10Dq$ problem in terms of orbital energies, e_i: It produces different expressions according as the spin orbitals $\psi_{t_{2g}}$ or ψ_{e_g} resulting from the splitting of the nd level are occupied or empty or are associated to spin-up or spin-down functions in a given configuration (Watson and Freeman, 1964). To avoid any confusion the simplest way is to start with orbital energies defined in the frame of a certain SCF theory and subtract from them all the interaction integrals J_{ij} and K_{ij} in order to have the difference $(I_{e_g} - I_{t_{2g}})$ directly (O'Donnell Offenhartz, 1969). If, however, the orbital energies are defined in conformity with the expression of the eigenvalues of the Fock–Roothaan Hamiltonian corresponding to an open shell, the energies assigned to the t_{2g} and e_g levels may vary with each spectroscopic state derived from the same configuration $(nl)^q$, and this is not very consistent with the basic concepts of the CLF theories. A more reasonable solution would be to take the open-shell Fock–Roothaan Hamiltonian corresponding to the average energy of the configuration as soon as the nl shell is not fully filled ($q < q_{\max} = 4l + 2$): According to Slater (1960) the average electronic interaction in an open shell is obtained by summing the contributions due to a given pair of spin orbitals ψ_i and ψ_j in each equivalent determinant and averaging them over all the possible determinants. To obtain the parameters f, a, and b of the standard open-shell theory, one has simply to identify the preceding result to the Roothaan formula *for spin orbitals*:

$$\frac{q(q-1)}{q_{\max}(q_{\max}-1)}\sum(J_{ij}-K_{ij}) = f^2\sum(aJ_{ij}-bK_{ij})$$

where $f = q/q_{\max}$ is the fractional occupation number of the nl shell. From the numerical values of the parameters f, a, and b, it is possible to derive unique expressions for the orbital energies $e_{t_{2g}}$ (O'Donnell Offenhartz, 1970). Semiempirical calculations carried out along these lines gave satisfactory results for the energy gap $10Dq$ in hexafluoride anions.

A number of explicit MO calculations have been performed with the aim of testing the validity of the CLF theories on selected compounds. The most popular system for such studies has been the complex $[NiF_6]^{4-}$ whose crystal-field theory describes as an ion $Ni^{2+}(d^8)$ surrounded by six negative charges $-e$ on the six F^- ions, giving a splitting $10Dq$ equal to 1380 cm^{-1} (exp.: 7250 cm^{-1}). The first quantum-mechanical attempt to refine the electrostatic model was most disappointing: If the point charges are replaced by spatial charge distributions built from F^- wavefunctions, the quantity $10Dq$ becomes negative; in other words, the t_{2g} and e_g levels are inverted (Kleiner, 1952). The correct ordering is slowly recovered as the overlap and covalence effect due to the ligands are introduced, but the simplicity of the primitive model disappears bit by bit. Actually, if the ionic description of the CLF theories is transposed into the MO method in the form of trial eigenvectors, the usual arrangement is reobtained with an orthogonalized orbital φ_- taking into account the covalence effects due to the "metal electrons" of the e_g antibonding level (Sugano and Shulman, 1963), but those covalence effects which can be associated with unbalanced ligand electrons" on the e_g bonding level seem to be still more important (Watson and Freeman, 1964), at least in a scheme without symmetry and equivalence restrictions on the MO's of the open shell. Finally, a realistic picture of the electron cloud must be achieved by the SCF method for the various states concerned in the $10Dq$ problem (Richardson *et al.*, 1969; Gladney and Veillard, 1969; Moskowitz *et al.*, 1970; Kleinman and Karplus, 1971), and the nonnegligible effects due to electron correlation (Hubbard *et al.*, 1966) have to be considered.

The evaluation of the quantity $10Dq$ is very sensitive to the approximations made when constructing effective Hamiltonians and computing basis integrals. In fact, very different values have been reported for the cluster $[NiF_6]^{4-}$ itself; the best *ab initio* treatment to date gives 4000 cm^{-1} for a calculation of $10Dq$ with orthogonalized ionic MO's and 5500 cm^{-1} with optimized SCF MO's (Wachters and Nieuwpoort, 1971). The agreement with experiment is fair, but it should be added that the calculation of a true transition energy by *ab initio* methods has almost nothing to do with the definition of an effective splitting parameter in a phenomenological theory, especially if the *ab initio* treatment takes the rearrangement of the electron cloud and the change in correlation energy into account, as this should be the case in the end.

The present situation has been depicted as a complete breakdown of the crystal-field model for transition-metal complexes (Jørgensen, 1971a). However, as concerns the question of excitation energies, the ancient and modern ways of thinking are somewhat reconciled if one adopts the concept of "*transition state*" recently developed by Slater (1972), that is to say a state where the occupation number of the orbitals are halfway between the initial and the final state of the excitation process. Using the X_α approximation, it can be shown that the transition energies can be expressed in terms of the orbital energies directly involved by the excitation in this fictitious transition state. The first applications of the X_α method to coordination compounds gave satisfactory results (Johnson and Wahlgreen, 1972).

C. Directed Valence Theory and Molecular Orbital Theory

In the various applications of the directed valence theory, attention concentrates on stereochemistry problems. The main features of the geometrical structure of a molecule are related to the hybridization states of its characteristic atoms, namely the central atom in the case of a transition-metal complex. The hybridization theory introduces an oriented element into the basic concept of *valence state* expressing the idea that the electrons of a given atom in a molecule have their spins random relative to each other if they are paired in chemical bonds with electrons coming from the surrounding atoms. In this process, a basis transformation is made at the atomic level by combining together the valence-shell orbitals of each atom so that the new basis functions have strongly directed probability distributions in coincidence—as much as possible—with the chemical bonds.

Such a transformation seems to be of little interest both from the interpretative and computational points of view: First, it gives no information on molecular conformations since the geometry of the molecule is an essential part of the imput data when calculating hybrids, and second, it cannot change the fundamental aspect of an *ab initio* calculation method nor its numerical results except possibly for immaterial factors on wavefunctions. As concerns the MO method itself, the mixing possibilities of an atomic set *s*, *p*, *d*, etc., inside canonical MO's transforming as basis vectors of the irreducible representations of the molecular symmetry group are not affected by any type of linear transformations. The name of hybridization could be reserved for those combinations which are missing in the symmetry-adapted MO's, but it is not more possible to preserve a sharp distinction between hybridization and symmetry-allowed mixing as soon as one considers molecules with few symmetry elements or extended basis sets (Dunn, 1970).

Actually, the classical directed-valence theory was not intended to facilitate the application of any quantum-mechanical calculation method, but rather to provide a basis set permitting the description of molecules in the frame of the *perfect pairing* approximation (see Penney, 1935). Because the hybrid orbitals of a given atom project in well-defined directions of the space, they are able to form strong bond orbitals with the hybrids of the nearest neighbors, each pair of chemically linked atoms being described by a two-center two-electron wavefunction (usually a bicentric MO with two electrons of opposite spins encompassing the metal and the nearest atom of each ligand or, alternatively, a Heitler–London localized electron pair). In its primitive form, the directed valence theory states that the arrangement of the ligands is determined by the directions of the hybrids available for σ bonds on the central atom. Furthermore, the orbitals remaining on the valence shell of the metal can combine with those of ligands and form π bonds reinforcing the stability of the complex or can be used in nonbonding electron pairs. Thus, the delocalized picture given by the symmetry-adapted MO's of Table II is replaced by a localized one where the hybridization of the central atom plays the main role: the σ bonds M–X are made by d^2sp^3 hybrids in octahedral compounds (e.g.,

$[Fe(CN)_6]^{4-}$), by dsp^2 hybrids in square planar compounds (e.g., $[Ni(CN)_4]^{2-}$) and by sp^3- or d^3s-hybrids in tetrahedral compounds (e.g., $[Cu(CN)_4]^{3-}$). Strong π bonds can be formed by the other three d orbitals in the first case, by two remaining d orbitals and one p orbital in the second case, and by three d orbitals in the third case (see Kimball, 1940).

Practically, the number and the nature of the orbitals to be hybridized are chosen in relation to the physical properties of each compound. In this context, the most useful information arises from magnetic susceptibility measurements giving the number of unpaired electrons unavailable for chemical bonding. Because the orbital angular momentum in a transition-metal complex is considered as virtually quenched, a magnetic moment μ different from zero means that there is a number n of unpaired electrons in the d shell of the metal ion given by the formula.

$$\mu = g_e[S(S+1)]^{1/2} \simeq [n(n+2)]^{1/2} \quad \text{a.u.},$$

where g_e is the Landé factor of the electron and $S = n/2$ the resultant spin of the system. For instance, $\mu = 0$ and $n = 0$ in the diamagnetic complexes of Fe^{2+}; consequently, the $3d$ orbitals can participate in the chemical bond, hence the d^2sp^3 hybridization typical of octahedral compounds. The same hybridization state can be retained in the corresponding paramagnetic complexes of Fe^{3+} where $\mu \simeq 3$ or $n = 1$, provided that the missing electron is supposed to belong originally to a $3d$ nonbonding lone pair. On the other hand, it is impossible to argue in an analogous way for the complexes of Fe^{2+} and Fe^{3+} with high magnetic moments because the metal has too many unpaired electrons and not enough $3d$ orbitals left for σ bonding. Similarly, the simple d^2sp^3-hybridization scheme breaks down if the system seems to have a too great number of electrons with respect to the number of distinct $3d$ orbitals available for its description; it is the case of the cluster $[NiF_6]^{4-}$ where there are two additional unpaired electrons above the subshell $3d^6$. In the classical valence theory, the transition metal complexes fall into two categories; those with low-spin values also called *inner* or *covalent* complexes, and those with high-spin values also called *outer* or *ionic* complexes. To save the familiar version of hybridization in the complexes of the second type, it is stated that the metal may form d^2sp^3 hybrids with $4d$ orbitals and even use the orbitals of higher shells for the extra electrons. Such an extension of the atomic basis set is not convincing because it is possible to describe the complexes of that sort in the frame of the MO method with basis functions including the orbitals $3d$, $4s$, and $4p$ only (Jørgensen, 1971b). Consequently, it should be possible to construct localized electron pairs of the usual type for these compounds by applying appropriate unitary transformations to the symmetry-adapted MO's (see below).

The underlying assumption in using the directed valence theory for stereochemical purposes is that the electron-pair wavefunctions built from the hybrids of the various atoms give a total energy corresponding to the most stable conformation effectively. However, qualitative arguments rather than explicit calculations are advanced to predict the best geometry. Some rely on the Pauling–Slater *criterion of maximum overlap*, and some on the Gillespie–Nyholm criterion of *maximum electron-pair repulsion*. The first point of view expresses the idea that the molecule reaches its maximum stability when the hybrids of each bond point toward each other as far as possible, because then the overlap integrals determining the bond strengths are maximum. The second point of view states that the stability is maximum for an arrangement of the various electron pairs maximizing their average distance apart, because then the repulsion energy of the electron system is minimum. Although the two criteria lead to the same conclusions in the case of perfect arrangements O_h or T_d, the second one offers the advantage of establishing a fruitful distinction between the repulsions due to lone-electron pairs and those due to bond-electron pairs (lone–pair, lone–pair > lone–pair, bond–pair > bond–pair, bond–pair). Nevertheless, if one recognizes the possibility of bent bonds in the directed valence theory, the maximum overlap principle turns out as flexible as the maximum electron-pair repulsion principle, since the direction of hybrids maximizing the overlap in a molecule varies according to the weight assigned to the lone pairs (Del Re *et al.*, 1966). Thus, the qualitative theories based on these two criteria are practically isomorphic, the one being only preferred for its anteriority, the other for its visualization power (see Berthier, 1971).

Finally, the picture of a molecule in terms of delocalized or localized one-electron functions, as it is done in the ligand-field theory and the directed-valency theory, can be considered as a result of equivalence transformations in the space spanned by a common atomic basis set. The two treatments only differ in the nature and size of the subspace in which they work: The directed-valence theory is primarily concerned with the description of the chemical bonds, σ and π, in the ground state; the ligand-field theory uses the d shell of the central atom to study the ground and excited states simultaneously. However, nothing requires taking delocalized one-electron functions in describing excited states, as soon as virtual localized orbitals can be constructed. In fact, computational processes have been developed in order to transform the delocalized spin orbitals of a

Slater determinant into quasi-localized ones without changing the values of the various observables: total density matrix, total energy, etc. (see Foster and Boys, 1960; Edmiston and Ruedenberg, 1963). The MO's obtained in this way for the ground state look very much like one-center and two-center hybrid orbitals which are intuitively associated to the various lone- and bond-electron pairs of a chemical formula, in spite of the presence of "tails" mainly due to the orthogonality constraints. A set of virtual antibonding MO's may be also relocalized, at least as concerns the ones which are the counterparts of the bonding MO's of the ground state. In so far as the form of these quasi-localized orbitals is rather insensitive to the choice of a particular localization process, one is justified in supposing that they result from the Edmiston–Ruedenberg maximum self-interaction criterion, and this can be considered as a molecular orbital translation of the maximum electron-pair repulsion principle since the maximization of the self-interaction energy terms implies the minimization of the interorbital Coulomb and exchange interaction terms (see England *et al.*, 1971). Conversely, if one starts with a localized electron-pair description, for instance derived from a ball-and-stick model obeying the prescriptions of the maximum electron-pair repulsion principle, it is possible to find the symmetry properties of the corresponding delocalized MO's in fair agreement with the predictions based on the classical Walsh rules for occupied orbitals (Thomson, 1968; Allen, 1972). This argument completes the proof of the equivalence of the different models of theoretical chemistry and suggests that the quality of the results obtained by the various methods of quantum chemistry depends on the quality of the atomic basis sets rather than on the methods themselves, provided that no simplifying assumption is introduced in course of application.

D. Nature of Coordination Bonds

It would be appropriate to conclude with some comments on the nature of the chemical bonding in coordination compounds. Unfortunately, the calculation methods of quantum chemistry do not give any direct answer to this question because they treat the various sorts of bonds in a chemical structure in like manner, the covalent bonds as well as the ionic ones. Moreover, the total value of an observable found after a quantum-mechanical computation, for instance the binding energy or the dipole moment of a molecule, is the resultant of too many intricate factors to supply immediate conclusions. It remains to be seen whether an analysis of the electronic structure in terms of atom and bond "*indices*" reveals some detail which might be correlated with the presence of a dative bond.

In fact, there are interesting peculiarities in the LCAO–SCF MO wave functions of organic and inorganic molecules with dative bonds. The overlap population $p_{D\to A}$ between two atoms belonging to electron-donor and electron-acceptor groups is very low, as if the formation of such chemical links does not require any bonding contribution. First observed in the treatment of the N → O bond as a two-electron two-center problem, where the doubly occupied MO was found even antibonding (Valdemoro, 1964), this remark still holds in *ab initio* calculations of the bonds N → O and N → B in amine–oxides and amine–boranes (Millié and Berthier, 1969), but not for phosphine–oxides (Demuynck and Veillard, 1970b), where PO is often written as a double bond. A typical example of a transition-metal complex with a low overlap population is the coordination-compound model $NH_3 \to ScH_3$ where $p_{N\to Sc} = 0.103$, as compared with the value $p_{F-Ti} = 0.378$ obtained for the molecule $FTiH_3$ (Stevenson and Lipscomb, 1969).

The formation of coordination compounds produces charge transfers δ in the direction of the dative bonds $D^+ \to A^-$, but a detailed analysis of the gross atomic populations shows that the electron transfers involve the donor and acceptor groups as a whole rather than the atoms directly linked in the classical chemical formula. In the case of the preceding coordination model, there is 0.12 e transferred from the donor NH_3 to the acceptor ScH_3. Higher charges are found for amine–oxides and amine–boranes: In a study of borazane $NH_3 \to BH_3$ using quasi-localized MO's (Aslangul, *et al.*, 1971), it was observed that the "*dative loge*" N → B is established between two groups of bonds with a total population very close to $+2$ for the donor and 0 for the acceptor. This gives a convenient criterion for dative bonds, the corresponding values being $+1$ on either side of a covalent bond. It should be also stated that the present descriptions based on the theoretical results available to date are compatible with the usual picture of coordination compounds, but they should not be accepted without some reserves, as long as the possible effects due to the basis set and the electron correlation are not fully explored.

ACKNOWLEDGMENTS

The author is indebted to Professor R. Moccia and Professor E. Scrocco, organizers of the Italian Easter School of Quantum Chemistry (Pisa, 1971), for giving him the impetus necessary to write this report on coordination compounds. He would like to express his gratitude to his former and present collaborators in this field for many constructive contributions and helpful critical comments.

REFERENCES

ALLEN, G. C., and CLACK, D. W. (1970). *J. Chem. Soc. A* p. 2668.

ALLEN, L. C. (1972). *Theor. Chim. Acta* **24**, 117.

ALLEN, T. L. (1957). *J. Chem. Phys.* **27**, 810.

ALTMANN, S. F. (1962). *In* "Quantum Theory" (R. D. Bates, ed.), Part II, p. 87. Academic Press, New York.

ASLANGUL, C., VEILLARD, A., DAUDEL, R., and GALLAIS, F. (1971). *Theor. Chim. Acta* **23**, 211.

ASTIER, M. (1970a). *C. R. Acad. Sci., Ser. C* **270**, 911.

ASTIER, M. (1970b). *Mol. Phys.* **19**, 209.

ASTIER, M., BERTHIER, G., and MILLIÉ, P. (1972). *J. Chem. Phys.* **57**, 5008.

BAIRD, N. C., SICHEL, J. M., and WHITEHEAD, M. A. (1968). *Theor. Chim. Acta* **11**, 38.

BALLHAUSEN, C. J. (1962). "Introduction to Ligand-Field Theory." McGraw-Hill, New York.

BALLHAUSEN, C. J. (1971). *Int. J. Quantum Chem., Symp.* **5**, 373.

BALLHAUSEN, C. J., and ANCMON, A. M. (1958). *Kgl. Dan. Vidensk. Selsk. Mat.-Fys. Medd.* **31**(9).

BALLHAUSEN, C. J., and GRAY, H. B. (1962). *Inorg. Chem.* **1**, 111.

BALLHAUSEN, C. J., and GRAY, H. B. (1964). "Molecular Orbital Theory." Benjamin, New York.

BASCH, H., VISTE, A., and GRAY, H. B. (1966a). *Theor. Chim. Acta* **3**, 458.

BASCH, H., VISTE, A., and GRAY, H. B. (1966b). *J. Chem. Phys.* **44**, 10.

BASCH, H., HOLLISTER, C., and MOSKOWITZ, J. W. (1970). *In* "Sigma MO Theory" (O. Sinanoğlu and K. B. Wiberg, eds.), p. 449. Yale Univ. Press, New Haven, Connecticut.

BERTHIER, G. (1964). *In* "Molecular Orbitals in Chemistry, Physics, and Biology" (P.-O. Löwdin and B. Pullman, eds.), p. 57. Academic Press, New York.

BERTHIER, G. (1971). *Colloq. Int. Cent. Nat. Rech. Sci.* **195**, 49.

BERTHIER, G., MILLIÉ, P., and VEILLARD, A. (1965). *J. Chim. Phys.* **62**, 8.

BERTHIER, G., DEL RE, G., and VEILLARD, A. (1966). *Nuovo Cimento, X*, **44**, 315.

BETHE, H. (1929). *Ann. Phys. (Leipzig)*, **3**, 135.

BROWN, R. D., and BURTON, P. G. (1970). *Theor. Chim. Acta* **18**, 309.

BROWN, R. D., and ROBY, K. R. (1970). *Theor. Chim. Acta.* **16**, 175, 194, 278, 291.

CALDER, V., MANN, D. E., SESHADRI, K. S., ALLAVENA, M., and WHITE, D. (1969). *J. Chem. Phys.* **51**, 2093.

CANADINE, R. M., and HILLIER, I. H. (1969). *J. Chem. Phys.* **50**, 2984.

CIULLIO, B., FURLANI, C., and SGAMELOTTI, A. (1971). *J. Chem. Soc. A* p. 1299.

CLACK, D. W., and FARRIMOND, M. S. (1971). *J. Chem. Soc. A* p. 299.

CONNOR, J. A., HILLIER, I. H., SAUNDERS, V. R., WOOD, M. H., and BARBER, M. (1972). *Mol. Phys.* **24**, 497.

COUTIÈRE, M. M., DEMUYNCK, J., and VEILLARD, A. (1972). *Theor. Chim. Acta* **27**, 281.

CUSACHS, L. C. (1965). *J. Chem. Phys.* **43**, Suppl., 157.

DACRE, P. D., and ELDER, M. (1971). *Chem. Phys. Lett.* **11**, 377.

DACRE, P. D., and ELDER, M. (1972). *J. Chem. Soc. Dalton* p. 1426.

DAHL, J. P., and BALLHAUSEN, C. J. (1968). *Advan. Quantum Chem.* **4**, 170.

DANON, J. (1963). *J. Chem. Phys.* **39**, 236.

DAUDEL, P., and DAUDEL, R. (1946). *J. Phys. Radium* **7**, 14.

DE BROUCKÈRE, G. (1967). *Bull. Soc. Chim. Belg.* **76**, 407.

DE BROUCKÈRE, G. (1970). *Theor. Chim. Acta* **19**, 310.

DEL RE, G., ESPOSITO, U., and CARPENTIERI, M. (1966). *Theor. Chim. Acta* **6**, 36.

DEMUYNCK, J., and VEILLARD, A. (1970a). *Chem. Phys. Lett.* **6**, 204.

DEMUYNCK, J., and VEILLARD, A. (1970b). *Chem. Commun.* p. 873.

DEMUYNCK, J., VEILLARD, A., and VINOT, G. (1971). *Chem. Phys. Lett.* **10**, 522.

DI SIPIO, L., TONDELLO, E., DE MICHELIS, G., and OLEARI, L. (1971). *Chem. Phys. Lett.* **11**, 287.

DUNN, T. M. (1970). *In* "Valency" (H. Eyring, ed.), Physical Chemistry, Vol. 5, p. 205. Academic Press, New York.

EDMISTON, C., and RUEDENBERG, K. (1963). *Rev. Mod. Phys.* **32**, 300.

ENGLAND, W., SALMON, L. S., and RUEDENBERG, K. (1971). *Fortschr. Chem. Forsch.* **23**, 31.

FENSKE, R. F., and RADTKE, D. D. (1968). *Inorg. Chem.* **7**, 479.

FENSKE, R. F., CAULTON, K. G., RADTKE, D. D., and SWEENEY, C. C. (1966). *Inorg. Chem.* **5**, 951.

FERREIRA, R. (1967). *Advan. Chem. Phys.* **13**, 55.

FISCHER, J., VEILLARD, A., and WEISS, R. (1972). *Theor. Chim. Acta* **24**, 317.

FOSTER, J. M., and BOYS, S. F. (1960). *Rev. Mod. Phys.* **32**, 300.

GALLAIS, F., VOIGT, D., and LABARRE, J. F. (1965). *J. Chim. Phys.* **62**, 761.

GALLAIS, F., DE LOTH, P., and LABARRE, J. F. (1966). *J. Chim. Phys.* **63**, 413.

GILLESPIE, R. J., and NYHOLM, R. S. (1957). *Quart. Rev. Chem. Soc.* **11**, 339.

GLADNEY, H. M., and VEILLARD, A. (1969). *Phys. Rev.* **180**, 385.

GOEPPERT-MAYER, M., and SKLAR, A. L. (1938). *J. Chem. Phys.* **6**, 645.

GRIFFITH, J. S. (1961). "The Theory of Transition-Metal Ions." Cambridge Univ. Press, London and New York.

GROPEN, C., and SEIP, H. M. (1971). *Chem. Phys. Lett.* **11**, 445.

HAISSINSKY, M. (1946). *J. Phys. Radium* **7**, 7.

HILLIER, I. H. (1969). *J. Chem. Soc. A* p. 878.

HILLIER, I. H. (1970). *J. Chem. Phys.* **52**, 1948.

HILLIER, I. H., and SAUNDERS, V. R. (1970). *Proc. Roy. Soc., Ser. A* **320**, 161.

HILLIER, I. H., and SAUNDERS, V. R. (1971). *Mol. Phys.* **22**, 1025.

HILLIER, I. H., and SAUNDERS, V. R. (1972). *Mol. Phys.* **23**, 449.

HINZE, J. (1968). *Forschr. Chem. Forsch.* **9**, 462.

HINZE, J., WHITEHEAD, M. A., and JAFFE, H. H. (1963). *J. Amer. Chem. Soc.* **85**, 184

HOFFMANN, R. (1963). *J. Chem. Phys.* **39**, 1397.

HOLLISTER, C., MOSKOWITZ, J. W., and BASCH, H. (1969). *Chem. Phys. Lett.* **3**, 185, 728.

HUBBARD, J., RIMMER, D. E., and HOPGOOD, F. R. A. (1966). *Proc. Phys. Soc., London* **88**, 13.

ICZKOWSKI, R. P., and MARGRAVE, J. L. (1961). *J. Amer. Chem. Soc.* **83**, 3547.

ILSE, F. E., and HARTMANN, H. (1951). *Z. Phys. Chem.* (*Leipzig*) **197**, 239.

JØRGENSEN, C. K. (1962). "Orbitals in Atoms and Molecules." Academic Press, New York.

JØRGENSEN, C. K. (1963). *Advan. Chem. Phys.* **5**, 33.

JØRGENSEN, C. K. (1969). "Oxidation Numbers and Oxidation States." Springer-Verlag, Berlin.

JØRGENSEN, C. K. (1971a). "Modern Aspects of the Ligand-Field Theory." North-Holland Publ., Amsterdam.

JØRGENSEN, C. K. (1971b). *Chimia* **25**, 109.

JØRGENSEN, C. K., PAPPALARDO, R., and SCHMIDTKE, H. H. (1963). *J. Chem. Phys.* **39**, 1422.

JØRGENSEN, C. K., HORNER, S. M., HATFIELD, W. E., and TYREE, S. Y. JR., (1967). *Int. J. Quantum Chem.* **1**, 191.

JOHANSEN, H. (1972). *Chem. Phys. Lett.* **17**, 569.

JOHNSON, K. H., and WAHLGREEN, U. (1972). *Int. J. Quantum. Chem., Symp.* **6**, 243.

KIBLER, M. R. (1970). *Chem. Phys. Lett.* **7**, 83.

KIDA, S., FUJITA, J., NAKAMOTO, K., and TSUCHIDA, R. (1957). *Bull. Chem. Soc. Jap.* **31**, 79.

KIMBALL, G. E. (1940). *J. Chem. Phys.* **8**, 188.

KLEIMAN, B., and KARPLUS, M. (1971). *Phys. Rev. B* **3**, 24.

KLEINER, W. H. (1952). *J. Chem. Phys.* **20**, 1784.

KLOPMAN, G. (1964). *J. Amer. Chem. Soc.* **86**, 1463, 4550.

KLOPMAN, G., and O'LEARY, B. (1970). *Forschr.* **15**, 445.

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

LÖWDIN, P.-O. (1964). *J. Mol. Spectrosc.* **14**, 112.

MCGARVEY, B. R. (1967). *J. Phys. Chem.* **71**, 51.

MANNE, R. (1966). *Theor. Chim. Acta* **6**, 299, 312.

MANNE, R. (1967). *J. Chem. Phys.* **46**, 4645.

MILLIÉ, P., and BERTHIER, G. (1969). *Colloq. Int. Cent. Nat. Rech. Sci.* **191**, 31.

MOSKOWITZ, J. W., HOLLISTER, C., HORNBACK, C. J., and BASCH, H. (1970). *J. Chem. Phys.* **53**, 2570.

MULLIKEN, R. S. (1949). *J. Chim. Phys.* **46**, 497.

MULLIKEN, R. S. (1955). *J. Chem. Phys.* **23**, 1833.

NESBET, R. K. (1955). *Proc. Roy. Soc. Ser. A* **230**, 312.

NEWTON, M. D., DE BOER, F. P., and LIPSCOMB, W. N. (1966). *J. Amer. Chem. Soc.* **88**, 2367.

NIEUWPOORT, W. C. (1965). *Philips Res. Rep.*, **20**, Suppl. 6.

O'DONNELL OFFENHARTZ, P. (1969). *J. Amer. Chem. Soc.* **91**, 5699.

O'DONNELL OFFENHARTZ, P. (1970). *J. Amer. Chem. Soc.* **92**, 2599.

OLEARI, L., DE MICHELIS, G., and DI SIPIO, L. (1966). *Mol. Phys.* **10**, 111.

ORGEL, L. E. (1955). *J. Chem. Phys.* **23**, 1004.

PAULING, L. (1948). *J. Chem. Soc.* p. 1461.

PENNEY, W. G. (1935). "The Quantum Theory of Valency." Methuen, London.

POPLE, J. A., SANTRY, D. P ., and SEGAL, G. A. (1965). *J. Chem. Phys.*, **43**, Suppl., 129,

RICHARDSON, J. W., VAUGHT, D. M., SOULES, T. F., and POWELL, R. P. (1969). *J. Chem. Phys.* **50**, 3633.

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

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

RUEDENBERG, K. (1951). *J. Chem. Phys.* **19**, 1433.

SANDERSON, R. T. (1952). *J. Amer. Chem. Soc.* **74**, 212, 4792.

SANDERSON, R. T. (1960). "Chemical Periodicity." Van Nostrand-Reinhold, Princeton, New Jersey.

SCHAEFFER, C. E. (1970). *Pure Appl. Chem.* **24**, 361.

SCHAEFFER, C. E., and JØRGENSEN, C. K. (1958). *J. Inorg. Nucl. Chem.* **8**, 143.

SCHAEFFER, C. E., and JØRGENSEN, C. K. (1965a). *Mol. Phys.* **9**, 401.

SCHAEFFER, C. E., and JØRGENSEN, C. K. (1965b). *Kgl. Dan. Vidensk. Selsk. Mat.-Fys. Medd.* **34**(13).

SCHLÄFER, H. L., and GLIEMAN, G. (1969). "Basic Principles of Ligand-Field Theory." Wiley, New York.

SCHLAPP, R., and PENNEY, W. G. (1932). *Phys. Rev.* **42**, 666.

SCHMIDTKE, H. H. (1971). *Theor. Chim. Acta* **20**, 92.

SENDER, M., and BERTHIER, G. (1958). *J. Chim. Phys.* **55**, 384.

SERRE, J. (1964). *In* "Molecular Orbitals for Chemistry, Physics, and Biology" (P.-O. Löwdin and B. Pullman, eds.), p. 133. Academic Press, New York.

SIDGWICK, N. V. (1923). *Trans. Faraday Soc.* **19**, 469.

SIDGWICK, N. V., and POWELL, H. M. (1940). *Proc. Roy. Soc., Ser. A* **176**, 153.

SLATER, J. C. (1960). "Quantum Theory of Atomic Structure," Vol. 1, Academic Press, New York.

SLATER, J. C. (1972). *Advan. Quantum Chem.* **6**, 1.

STEVENSON, P. E., and LIPSCOMB, W. N. (1969). *J. Chem. Phys.* **50**, 3306.

SUGANO, S., and SHULMAN, R. G. (1963). *Phys. Rev.* **130**, 517.

TANABE, Y., and SUGANO, S. (1954). *J. Phys. Soc. Jap.* **9**, 753, 766.

THOMSON, H. B. (1968). *Inorg. Chem.* **7**, 605.

TOSSELL, J. A., and LIPSCOMB, W. N. (1972). *J. Amer. Chem. Soc.* **94**, 1505.

TRAPPENIERS, N. J., DE BROUCKÈRE, G., and TEN SELDAM, C. A. (1971). *Chem. Phys. Lett.* **8**, 327.

TRAPPENIERS, N. J., DE BROUCKÈRE, G., and TEN SELDAM, C. A. (1972). Personal communication.

VALDEMORO, C. (1964). Thesis, Univ. of Madrid, Madrid.

VAN HOOYDONK, G., and EECKHAUF, Z. (1970). *Ber. Bunsenges. Phys. Chem.* **74**, 323, 327.

VAN VLECK, J. H. (1932). *Phys. Rev.* **41**, 208.

VAN VLECK, J. H. (1935). *J. Chem. Phys.* **3**, 803.

VAN VLECK, J. H. (1939). *J. Chem. Phys.* **7**, 61, 72.

VEILLARD, A. (1969). *Chem. Commun.* pp. 1022, 1427.

VEILLARD, A., and PULLMAN, B. (1965). *J. Theor. Biol.* **8**, 307, 317.

WACHTERS, A. J. H., and NIEUWPOORT, W. C. (1971). *Int. J. Quantum Chem., Symp.*, **5**, 391.

WATSON, R. E., and FREEMAN, A. J. (1964). *Phys. Rev. A* **134**, 1526.

WOLFSBERG, M., and HELMHOLZ, L. (1952). *J. Chem. Phys.* **20**, 837.

YERANOS, W. A., and HASMAN, D. A. (1967). *Z. Naturforsch. A* **22**, 170.

Effective Hamiltonian Methods for Molecular Collisions

DAVID A. MICHA
Quantum Theory Project
Departments of Chemistry and Physics
University of Florida
Gainesville, Florida

I. Introduction

A large number of atomic and molecular properties are observed in processes that involve collision events. Even in the relatively few cases where interaction forces are well known (e.g., when electronic and nuclear motions may be separated) there remains a need for knowledge of cross sections for the collision events. We present here an operator approach of general applicability to molecular collisions. Generality is important within the field under consideration because of the great variety of collision partners, such as electrons, atoms, diatomics or polyatomics, and their ions, that may be encountered. The operator approach is used to provide a unified treatment of two wide areas, namely scattering resonances and optical potentials. These two areas, although frequently discussed separately, have a common origin when brought up in terms of effective hamiltonians. The present treatment is being restricted to direct processes, i.e., it concerns both elastic and inelastic processes but not reactive ones. This is due partly to limitations in space and partly to lack at present of satisfactory treatments of resonance and optical potentials for reactive collisions. Although the main thrust of this work is on the formalism, we introduce at the beginning an example of collision between an atom and a diatom and return to it several times to illustrate the subjects under consideration.

The purpose of Section II is to introduce the reader to the formalism of collision theory (Lippmann and Schwinger, 1950; Gell-Mann and Goldberger, 1953) in its time-dependent and time-independent forms, to present the main definitions and concepts for subsequent use. It does also discuss the rate of change of observables (Lippmann, 1965) and in particular cross sections. More details on this may be found in the works by Roman (1965), Newton (1966), Goldberger and Watson (1964), and Levine (1969).

Section III discusses the partitioning technique, which has been extensively used since its introduction in collision theory (Feshbach, 1958a, 1962, 1967; Newton and Fonda, 1960; Fano, 1961) and in work on bound states (Riesenfeld and Watson, 1956; Löwdin, 1963, 1966). It is done here by defining diagonal and nondiagonal parts of operators, in a form suitable for partitioning of operator equations. Different choices of projection operators are illustrated for molecular collisions, and a detailed discussion is presented on upper and lower bounds to effective interactions (Kato, 1951; Bazely and Fox, 1961; Sugar and Blankenbecler, 1964; Lindner and Löwdin, 1968; Spruch, 1969; Miller, 1969; Weinhold, 1972).

Section IV presents distortion, and adiabatic and sudden approximations based on different splittings of effective hamiltonians. Of these, the adiabatic approximation has been extensively studied in connection with separation of electronic and nuclear motions in collisions (Bates, 1961; Nikitin, 1968; Levine, 1968; Pack and Hirschfelder, 1968, 1970; Levine *et al.*, 1969; Evans *et al.*, 1971; Browne, 1971; Kolos, 1970) but has seldom been applied to heavy particles. Alternatives to this approximation have been described by F. T. Smith (1969) and by O'Malley (1971). The sudden approximation has mainly been used for nuclear collisions (Chase, 1956 Barrett, 1964; Austern and Blair, 1965; Glendenning, 1969) under a different name but seems promising also in molecular problems. This section describes corrections to the usually employed lowest order version of these approximations.

Section V considers resonance energies and widths in the general case of overlapping resonances for couplings of arbitrary strength and concentrates on their determination as energy-independent quantities (Newton and Fonda, 1960; Micha and Brändas, 1971). It describes the time dependence of the corresponding decaying states (Newton,1966; Messiah, 1961) and discusses atom–diatom resonances within both the distortion and adiabatic approximations (Micha, 1967a,b, 1973; Levine *et al.*, 1968; Muckerman and Bernstein, 1970; Von Seggern and Toennies, 1969; Burke *et al.*, 1969; Levine, 1970).

A similar path is followed in Section VI on optical potentials, for which we first discuss the time-dependent interpretation and energy dependence (Watson, 1957; Fetter and Watson, 1965); we proceed to their approximations and then their application to energy transfer in atom–diatom collisions (Micha, 1969; Rotenberg, 1971; Ross and Greene, 1970).

The cited literature is detailed only insofar as it relates to molecular collisions. We refer to reviews for applications of similar effective hamiltonian methods to nuclear collisions (Feshbach, 1958b; Glendenning, 1969; Hodgson, 1971; also see Lane and Robson, 1966; Robson and Lane, 1967), to electron scattering by atoms (Burke, 1965; Smith, 1966; Taylor, 1970; Chen,1970), and to electron scattering by molecules (Bardsley and Mandl, 1968; Chen, 1969). Earlier, and important, work using wave–mechanical language is extensively described by Mott and Massey (1965).

Although much of the material here is well known or has been recently published, some of the results on the sudden approximation and on opticals potentials have not previously appeared in the literature.

II. Collision States and Transition Probabilities

A. Channel States and the Atom–Diatom Example

We consider collisions between two particles with internal structure A and M (atoms or molecules) which may in general change their internal state, as shown in Fig. 1 in the center-of-mass reference frame. The

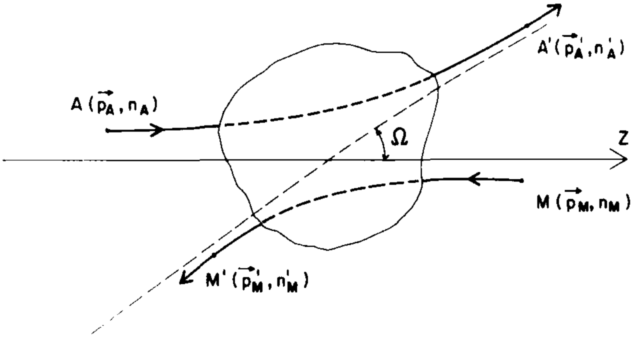


Fig. 1. Collision between particles A and M in the center-of-mass frame, where $\mathbf{p}_M = -\mathbf{p}_A$ and $\mathbf{p}_M' = -\mathbf{p}_A'$.

hamiltonians and eigenstates describing A and M in their own reference frame satisfy

$$H_A u_{n_A}^{(A)} = W_{n_A}^{(A)} u_{n_A}^{(A)} \tag{1a}$$

$$H_M u_{n_M}^{(M)} = W_{n_M}^{(M)} u_{n_M}^{(M)} \tag{1b}$$

where the internal energies $W^{(A)}$ or $W^{(M)}$ are taken as zero for the ground states. Introducing relative coordinates, the corresponding reduced mass m, and relative momentum $\mathbf{p}$, we write for the overall internal motion

$$H_I u_n = W_n u_n \tag{2}$$

with

$$H_I = H_A + H_M \tag{3}$$

$$u_n = u_{n_A}^{(A)} \cdot u_{n_M}^{(M)}, \tag{4}$$

$$W_n = W_{n_A}^{(A)} = W_{n_M}^{(M)}, \tag{5}$$

where n is the collection of quantum numbers of both A and M. A scattering channel is defined then by stating the nature of A and M and giving n. The total hamiltonian is

$$H = K + H_I + V \tag{6}$$

with K the relative kinetic energy operator and V the interaction between A and M. In the free-motion region where $V = 0$ we introduce plane wave states $\chi_\mathbf{p}$ which satisfy

$$K\chi_\mathbf{p} = (p^2/2m)\chi_\mathbf{p}, \tag{7a}$$

$$\langle \chi_\mathbf{p} | \chi_{\mathbf{p}'} \rangle = \delta(\mathbf{p} - \mathbf{p}') \qquad \text{and} \qquad \int d\mathbf{p}\, |\chi_\mathbf{p}\rangle\langle\chi_\mathbf{p}| = 1, \tag{7b}$$

and write for the state specified by $\nu = (\mathbf{p}, n)$

$$(E - H_0)\Psi_{0\nu} = 0 \tag{8}$$

where $E = p^2/2m + W_n$ and

$$\Psi_{0\nu} = \chi_{\mathbf{p}} u_n \tag{9}$$

is the channel state satisfying

$$\langle \Psi_{0\nu} | \Psi_{0\nu'} \rangle = \delta_{\nu\nu'}, \tag{10a}$$

$$\int d\nu |\Psi_{0\nu}\rangle\langle\Psi_{0\nu}| = 1. \tag{10b}$$

Here we integrate over directions and magnitude of $\mathbf{p}$ and sum over all quantum numbers n, including in principle also those n corresponding to dissociated A or M.

To illustrate the following developments we consider in greater detail the scattering of an atom A in the ^{1}S state by a diatomic molecule BC in the $^1\Sigma$ state as a result of which there may be rotational or vibrational transitions in the diatom, but no electronic transitions. This is a relatively simple case to describe formally, yet it contains the main features (e.g., energy transfer) present in more complex situations. The coordinates are shown in Fig. 2 with $\mathbf{r}$ giving the internal position and $\mathbf{R}$ the intermolecular one. We have in this case, for vibration–rotation quantum numbers (njm_j),

$$H_{\mathrm{I}} u_{njm_j}(\mathbf{r}) = W_{nj} u_{njm_j}(\mathbf{r}), \tag{11}$$

$$-(\hbar^2/2m)\nabla^2 \chi_{\mathbf{p}}(\mathbf{R}) = (p^2/2m)\chi_{\mathbf{p}}(\mathbf{R}), \tag{12}$$

with

$$\chi_{\mathbf{p}}(\mathbf{R}) = (2\pi\hbar)^{-3/2} \exp(i\mathbf{p}\mathbf{R}/\hbar) \tag{13}$$

and

$$V(R, r, \gamma) = \sum_{\kappa=0}^{\infty} v_\kappa(R, \rho) P_\kappa(\cos\gamma) \tag{14}$$

with $\rho = r - r_e$ the deviation from the equilibrium internal distance r_e. This interaction is highly repulsive at short distances and goes as R^{-6} at large distances.

In this case it is convenient to introduce conservation laws for the whole system, and in particular conservation of the total angular momentum

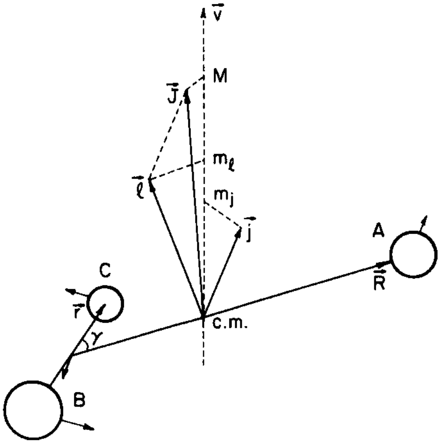


Fig. 2. Position vectors **R** and **r** for atom A and diatomic BC. The *z* axis has been chosen along the initial relative velocity **v**. Angular momenta, **j**, **l** and **J** and their projections are also shown.

$\mathbf{J} = \mathbf{l} + \mathbf{j}$, where **l** and **j** are orbital and rotational angular momenta, respectively, and of the total parity Π (Arthurs and Dalgarno, 1960; Takayanagi, 1965; Newton, 1966; Micha, 1967b). This may be done expanding first $\chi_{\mathbf{p}}(\mathbf{R})$ in spherical harmonics of the relative angles Ω and of the momentum angles ω_p, and then introducing states in the $(ljJM)$ representation

$$\mathscr{Y}^{M}_{Jnjl}(\Omega, \mathbf{r}) = \sum_{m_l m_j} \langle l m_l j m_j | JM \rangle i^l Y_l^{m_l}(\Omega) u_{njm_j}(\mathbf{r}), \tag{15}$$

where Clebsch-Gordan coefficients have been used (e.g., see Edmonds, 1960). Choosing u_{njm_j} with parity $(-1)^j$, the overall parity of $\mathscr{Y}$ in Eq. (15) is

$$\Pi = (-1)^{l+j}. \tag{16}$$

Channel states may be written in terms of these functions in the form

$$\Psi^{(0)}_{\mathbf{p}njm_j}(\mathbf{R}, \mathbf{r}) = (2\pi\hbar)^{-3/2} \sum_{JMl} 4\pi j_l(pR/\hbar)\, \mathscr{X}^{Mm_j}_{Jlj}(\omega_p)^* \mathscr{Y}^{M}_{Jnjl}(\Omega, \mathbf{r}), \tag{17}$$

where

$$\mathscr{X}^{Mm_j}_{Jlj}(\omega_p) = \langle l m_l j m_j | JM \rangle Y_l^{m_l}(\omega_p), \tag{18}$$

and $j_l(x)$ is the spherical Bessel function. The set of states in Eq. (15) may be used to expand the terms in H. Introducing the collective index $\lambda = (njl)$, the results are

$$K_{\lambda'\lambda} = -\frac{\hbar^2}{2m}\left[\frac{1}{R}\frac{d^2}{dR^2}R - \frac{l(l+1)}{R^2}\right]\delta_{\lambda'\lambda}, \tag{19a}$$

$$(H_\mathrm{I})_{\lambda'\lambda} = W_{nj}\,\delta_{\lambda'\lambda}, \tag{19b}$$

$$V^J_{\lambda'\lambda}(R) = \langle \mathscr{Y}^M_{J\lambda'}|V|\mathscr{Y}^M_{J\lambda}\rangle = \sum_{x=0}^{\infty} v^{(\kappa)}_{n'j',nj}(R)a_\kappa(l'j', lj; J). \tag{19c}$$

These matrix elements are independent of M and Π and diagonal in J, in accordance with the Wigner–Eckart theorem, because each of the hamiltonian terms are invariant under overall rotation and inversions. The coefficients in Eq. (19c) are

$$\begin{aligned} a_\kappa(l'j', lj; J) &= i^{l+j-l'-j'}(-1)^{j'+j+J}[(2l+1)(2j+1)(2l'+1)(2j'+1)]^{1/2} \\ &\quad\times \begin{Bmatrix} J & j' & l' \\ \kappa & j & l \end{Bmatrix} \begin{pmatrix} l & \kappa & l' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} j & \kappa & j' \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{20}$$

In matrix notation we can then write Eq. (6) for fixed J as

$$\mathbf{H}^J = \mathbf{K} + \mathbf{H}_\mathrm{I} + \mathbf{V}^J \tag{21}$$

and Eq. (8) as

$$[k^2 - (2m/\hbar^2)(R\mathbf{K}R^{-1} + \mathbf{H}_\mathrm{I})]\boldsymbol{\psi}_0(k; R) = 0 \tag{22a}$$

with the solution

$$\begin{aligned} [\boldsymbol{\psi}_0(R)]_{\lambda'\lambda} &= 2ik_\lambda^{1/2}Rj_l(k_\lambda R)\delta_{\lambda'\lambda} \\ &\sim k_\lambda^{-1/2}[e^{i(k_\lambda R - l\pi/2)} - e^{-i(k_\lambda R - l\pi/2)}]\delta_{\lambda'\lambda}, \end{aligned} \tag{22b}$$

where $k^2 = 2mE/\hbar^2$, $k_\lambda{}^2 = k^2 - 2mW_{nj}/\hbar^2$ and the normalization has been chosen so that asymptotically $\boldsymbol{\psi}_0$ contains incoming and outgoing spherical waves of fixed current density. Channel states may now be rewritten in the form

$$\begin{aligned} \Psi^{(0)}_{\mathbf{p}njm_j}(\mathbf{R}, \mathbf{r}) &= (2\pi\hbar)^{-3/2}\sum_{JMl}(-2\pi i)k_\lambda^{-1/2}\mathscr{X}^{Mm_j}_{Jlj}(\omega_p)^* \\ &\quad\times \sum_{\lambda'} R^{-1}[\psi_0(R)]_{\lambda'\lambda}\mathscr{Y}^M_{J\lambda'}(\Omega, \mathbf{r}). \end{aligned} \tag{23}$$

These matrix equations and their extensions in following sections are particularly convenient for computational purposes.

B. The Time-Dependent Treatment

A time-dependent description of stationary scattering processes is convenient for several reasons, although it is not necessary inasmuch as both the scattering equations and boundary conditions are independent of time. Among these reasons we mention (1) the possibility of constructing wavepackets for the collision partners, that describe more accurately than plane waves the physical situation, (2) the natural incorporation of the role of causality, whereby in an experiment with fixed initial conditions no scattered waves should be detected at times prior to arrival to the collision region, and (3) mathematical advantages that result from using normalized wave functions, such as the possibility of interchanging limiting procedures and integration.

We begin by writing for the wavepacket in the absence of interaction

$$\Psi_0(t) = \int dv\, c_v \Psi_{0v} e^{iEt/\hbar} \tag{24}$$

where c_v is obtained from the coefficients making up the wavepackets for A and M and satisfies

$$\int dv\, |c_v|^2 = 1. \tag{25}$$

In a center-of-mass frame it is a function of the difference $\mathbf{p} - \mathbf{p}_0$ between the relative momentum $\mathbf{p}$ and its average initial value $\mathbf{p}_0$, and of n_A and n_M. The time t running between plus and minus infinity may be chosen so that $t = 0$ corresponds to the instant at which the centers of mass of A and M would come closest if the interaction V were zero.

When the interaction V is included, the state $\Psi(t)$ is given by the solution of

$$(i\hbar\, \partial/\partial t - H)\Psi(t) = 0 \tag{26}$$

that satisfies $\langle \Psi(t) | \Psi(t) \rangle = 1$ and has the predetermined value $\Psi(t_0)$ at the initial time t_0. Alternatively the solution may be expressed in terms of the time-evolution operator $U(t, t_0)$ (e.g., see Löwdin, 1967) by

$$\Psi(t) = U(t, t_0)\Psi(t_0) \tag{27}$$

with

$$U(t, t_0) = \exp[-iH(t - t_0)/\hbar], \tag{28}$$

a function only of the elapsed time $t - t_0$.

We may now add the restrictions of causality by modifying the previous boundary conditions. We introduce the Green functions $G^{(\pm)}(t, t_0)$ that satisfy

$$(i\hbar\, \partial/\partial t - H)G^{(\pm)}(t, t_0) = \delta(t - t_0), \tag{29}$$

$$G^{(\pm)}(t, t_0) = 0 \qquad \text{for} \quad t \lessgtr t_0, \tag{30}$$

respectively. Here $G^{(+)}$ is the retarded Green function and $G^{(-)}$ the advanced one, and both have dimensions of action.

For $t \neq t_0$, Eq. (29) reduces to that for U, and it is immediately verified by differentiation with respect to t that the solutions are

$$G^{(\pm)}(t, t_0) = \mp \frac{i}{\hbar}\,\theta[\pm(t - t_0)]U(t - t_0), \tag{31}$$

where we have used the step function

$$\begin{aligned} \theta(\tau) &= 1 \qquad \text{for} \quad \tau > 0, \\ &= 0 \qquad \text{for} \quad \tau < 0, \end{aligned} \tag{32}$$

for which $d\theta/d\tau = \delta(\tau)$. The Green functions are seen to satisfy

$$G^{(\pm)}(\tau)^\dagger = G^{(\mp)}(-\tau). \tag{33}$$

Equation (27) may be rewritten, using them, as

$$\Psi(t) = \pm i\hbar G^{(\pm)}(t - t_0)\Psi(t_0), \qquad \text{for} \quad t \gtrless t_0. \tag{34}$$

Based on this expression we shall obtain an integral equation for $\Psi(t)$. We first define the unperturbed Green functions $G_0^{(\pm)}$ as the solutions of

$$(i\hbar\, \partial/\partial t - H_0)G_0^{(\pm)}(t - t_0) = \delta(t - t_0). \tag{35}$$

Next we multiply both sides of the equality

$$i\hbar\, \partial G^{(\pm)}(t' - t_0)/\partial t' = (H_0 + V)G^{(\pm)}(t' - t_0) + \delta(t' - t_0) \tag{36}$$

by $G_0^{(\pm)}(t - t')$ and integrate over all t'. Integrating by parts the left-hand side and using Eq. (35) the result is

$$G^{(\pm)}(t - t_0) = G_0^{(\pm)}(t - t_0) + \int_{-\infty}^{+\infty} dt'\, G_0^{(\pm)}(t - t')VG^{(\pm)}(t' - t_0). \tag{37}$$

Replacing this in Eq. (34) we get, restricting ourselves to times $t > t_0$,

$$\Psi(t) = i\hbar G_0^{(+)}(t - t_0)\Psi(t_0) + \int_{-\infty}^{+\infty} dt'\, G_0^{(+)}(t - t')V\Psi(t'), \tag{38}$$

where $\Psi(t)$ depends on t_0. In the limit $t_0 \to -\infty$ we define

$$\Psi_{\text{in}}(t) = \lim_{t_0 \to -\infty} i\hbar G_0^{(+)}(t - t_0)\Psi(t_0), \tag{39}$$

$$\Psi^{(+)}(t) = \lim_{t_0 \to -\infty} \Psi(t), \tag{40}$$

where Ψ_{in} is an incoming wavepacket prepared in the past from eigenstates of H and evolving thereon freely into the present, while $\Psi^{(+)}(t)$ is the actual state at t resulting from the scattering of Ψ_{in}. The basic assumption of scattering theory states that we can replace

$$\Psi_{\text{in}}(t) \Rightarrow \Psi_0^{(+)}(t) = \lim_{t_0 \to -\infty} i\hbar G_0^{(+)}(t - t_0)\Psi_0(t_0), \tag{41}$$

which requires knowledge only of the eigenstates of H_0. The resulting integral equation for scattering is then

$$\Psi^{(+)}(t) = \Psi_0^{(+)}(t) + \int_{-\infty}^{+\infty} dt'\, G_0^{(+)}(t - t')V\Psi^{(+)}(t'). \tag{42}$$

The basic assumption is in practice justified provided the collision region occupies only a finite portion of space. In this case, at sufficiently early times the wavepackets of A and M move in regions where $V = 0$. A counterexample may be found in quantum electrodynamics, where photons and the quantized electron field interact all over space, and renormalization procedures become necessary.

Using the advanced Green functions, we would have arrived at Eq. (42) but with $(-)$ upper indices after introducing outgoing wavepackets

$$\Psi_{\text{out}}(t) = \lim_{t_0 \to +\infty} i\hbar G^{(-)}(t - t_0)\Psi(t_0) \tag{43}$$

and making our basic assumption.

C. The Time-Independent Treatment

The previous treatment may be recast into a time-independent context that is more closely related to the standard collision theory. This is best done introducing the nonsymmetrical Fourier transform in time

$$f(s) = \int_{-\infty}^{+\infty} dt\, e^{ist/\hbar} f(t), \tag{44}$$

$$f(t) = \int_{-\infty}^{+\infty} \frac{ds}{2\pi\hbar}\, e^{-ist/\hbar} f(s), \tag{45}$$

where s has units of energy and we use the same symbol to signify a function and its transform, distinguishing them by their argument.

We can immediately obtain the transform of Eq. (42) making use of the well-known relationship

$$\int_{-\infty}^{+\infty} dt\, e^{ist/\hbar} \int_{-\infty}^{+\infty} dt'\, f(t-t')g(t') = f(s)g(s) \tag{46}$$

which gives

$$\Psi^{(+)}(s) = \Psi_0^{(+)}(s) + G_0^{(+)}(s)V\Psi^{(+)}(s), \tag{47}$$

where

$$\begin{aligned} G_0^{(+)}(s) &= -(i/\hbar)\int_{-\infty}^{+\infty} dt\, e^{ist/\hbar}\theta(t)e^{-iH_0t/\hbar} \\ &= \lim_{\varepsilon\to 0+}(-i/\hbar)\int_0^{\infty} dt\, e^{i(s+i\varepsilon-H_0)t/\hbar} \end{aligned} \tag{48}$$

after using the definition of $\theta(t)$ and interchanging limit and integration procedures. This interchange is justified provided we have, after operating with $G_0^{(+)}$ on a function, a well-behaved integrand. This is the case here, since we are working with normalized wavepackets. Carrying out the integration we get

$$G_0^{(+)}(s) = \lim_{\varepsilon\to 0+}(s+i\varepsilon-H_0)^{-1} = (s^{(+)}-H_0)^{-1}, \tag{49}$$

where $s^{(+)} = s + i\varepsilon$ and the limit is implicit and must be taken only after scalar products involving $G_0^{(+)}$ and normalizable states have been performed. Equation (49) gives the free propagator for our collision process.

In this way we obtain the Lippmann–Schwinger equation for colliding wavepackets,

$$\Psi^{(+)}(s) = \Psi_0^{(+)}(s) + G_0^{(+)}(s)V\Psi^{(+)}(s). \tag{50}$$

We stress that the states appearing here are transforms of wavepackets and not solutions of the time-independent Schrödinger equation, which is the reason why we use the variable s for their arguments.

Similarly to Eq. (49) we define

$$G^{(+)}(s) = (s^{(+)} - H)^{-1}, \tag{51}$$

and using

$$(A-B)^{-1} = A^{-1} + A^{-1}B(A-B)^{-1} = A^{-1} + (A-B)^{-1}BA^{-1}, \tag{52}$$

we obtain

$$G^{(+)}(s) = G_0^{(+)}(s) + G_0^{(+)}(s)VG^{(+)}(s) \tag{53}$$

$$= G_0^{(+)}(s) + G^{(+)}(s)VG_0^{(+)}(s). \tag{54}$$

Multiplying both sides of Eq. (50) by $(s^{(+)} - H_0)$ we find that

$$\Psi^{(+)}(s) = W^{(+)}(s)\Psi_0^{(+)}(s) \tag{55}$$

with

$$\begin{aligned} W^{(+)}(s) &= G^{(+)}(s)(s^{(+)} - H_0) \\ &= 1 + G^{(+)}(s)V, \end{aligned} \tag{56}$$

the wave operator for the scattering problem. Replacing Eq. (55) in the right-hand side of Eq. (50),

$$\Psi^{(+)}(s) = \Psi_0^{(+)}(s) + G_0^{(+)}(s)T^{(+)}(s)\Psi_0^{(+)}(s) \tag{57}$$

where we have introduced the transition operator

$$T^{(+)}(s) = VW^{(+)}(s) = V + VG^{(+)}(s)V. \tag{58}$$

Similar equations are obtained in terms of advanced-type propagators.

The connection with stationary scattering states may be done returning to the free wavepacket in Eq. (24) and taking its Fourier transform, which gives

$$\Psi_0^{(+)}(s) = \int ds\, 2\pi\hbar\, \delta(s - E) \int d\nu\, c_\nu \Psi_{0\nu} \tag{59}$$

with $\delta(s)$ the transform of unity. Operating on both sides with $W^{(+)}(s)$ and taking the antitransform of the result $\Psi^{(+)}(s)$, we get

$$\Psi^{(+)}(t) = \int d\nu\, c_\nu \Psi_\nu^{(+)} e^{-iEt/\hbar}, \tag{60a}$$

$$\Psi_\nu^{(+)} = W_\nu^{(+)}(E)\Psi_{0\nu}. \tag{60b}$$

Comparing Eq. (60a) with Eq. (24) we find that $\Psi_\nu^{(+)}$ is the stationary scattering state of interest. Equation (60b) may be written in the form

$$(E^{(+)} - H)\Psi_\nu^{(+)} = i\varepsilon\Psi_{0\nu}, \tag{61}$$

which replaces the Schrödinger equation for bound states, or in the form

$$\Psi_\nu^{(+)} = i\varepsilon G^{(+)}(E)\Psi_{0\nu}, \tag{62}$$

which shows that the wave operator may in this case be defined only in terms of H and is independent of any splitting in unperturbed part plus perturbation. The normalization of $\Psi_v^{(+)}$ follows from Eq. (62) with the help of the relation

$$G^{(+)}(E)^{\dagger}G^{(+)}(E') = (E - E' - 2i\varepsilon)^{-1}[G^{(+)}(E') - G^{(+)}(E)^{\dagger}], \tag{63}$$

and of Eq. (53), which lead to

$$\langle \Psi_v^{(+)} | \Psi_{v'}^{(+)} \rangle = \langle \Psi_{0v} | \Psi_{0v'} \rangle = \delta_{vv'} \tag{64}$$

in the limit $\varepsilon \to 0$, for fixed E and E'. These stationary scattering states together with the set of orthonormalized bound states Ψ_b form a complete set. Hence Eq. (10b) is replaced by

$$\sum_b |\Psi_b\rangle\langle\Psi_b| + \int dv\, |\Psi_v^{(+)}\rangle\langle\Psi_v^{(+)}| = 1. \tag{65}$$

Similar relations could be given for $\Psi_v^{(-)}$.

The wave-mechanical description of scattering is recovered working in the coordinate representation. We then write, for a normalizable state $\xi(\mathbf{R}, \mathbf{r})$,

$$G_0^{(+)}\xi(\mathbf{R}, \mathbf{r}) = \int d\mathbf{R}'\, d\mathbf{r}'\, G_0^{(+)}(\mathbf{R}\mathbf{r}, \mathbf{R}'\mathbf{r}')\xi(\mathbf{R}'\mathbf{r}'), \tag{66a}$$

$$G_0^{(+)}(\mathbf{R}\mathbf{r}, \mathbf{R}'\mathbf{r}') = \sum_n u_n(\mathbf{r})^* G_{0n}^{(+)}(\mathbf{R}, \mathbf{R}')u_n(\mathbf{r}'), \tag{66b}$$

$$\begin{aligned} G_{0n}^{(+)}(\mathbf{R}, \mathbf{R}') &= \langle \mathbf{R} | (E^{(+)} - W_n - K)^{-1} | \mathbf{R}' \rangle \\ &= -(4\pi)^{-1}(2m/\hbar^2)e^{i\mathbf{k}_n|\mathbf{R}-\mathbf{R}'|}/|\mathbf{R} - \mathbf{R}'|, \end{aligned} \tag{66c}$$

where the last line follows from integration in the complex k-plane, and $k_n{}^2 = (2m/\hbar^2)(E - W_n)$, with $\operatorname{Im} k_n \geq 0$. For large R, and using the asymptotic form of Eq. (66c), one finds from Eq. (57) for a stationary state the familiar expansion

$$\Psi_v^{(+)}(\mathbf{R}, \mathbf{r}) \underset{R \to \infty}{\sim} (2\pi\hbar)^{-3/2} \sum_{n'} u_{n'}(\mathbf{r})[e^{i\mathbf{k}_n \cdot \mathbf{R}}\delta_{nn'} + R^{-1} \exp(ik_{n'}R) f_{n'n}(\Omega)] \tag{67a}$$

with

$$f_{n'n}(\Omega) = -(4\pi)^{-1}(2m/\hbar^2)(2\pi\hbar)^3\langle \Psi_{0v'} | T^{(+)}(E) | \Psi_{0v} \rangle, \tag{67b}$$

where now $v' = (\hbar k_{n'} \mathbf{R}/R, n')$.

For our atom–diatom problem, Eq. (66b) may further be rearranged as an expansion in the $\mathscr{Y}$-set, which gives

$$G_0^{(+)}(\mathbf{R}\mathbf{r}, \mathbf{R}'\mathbf{r}') = \sum_{JM\lambda} \mathscr{Y}_{J\lambda}^{M}(\Omega, \mathbf{r})^* g_{\lambda,0}^{(+)}(R, R') \mathscr{Y}_{J\lambda}^{M}(\Omega', \mathbf{r}'). \tag{68}$$

The reduced radial Green functions $g_{\lambda,0}^{(+)}$ are elements of the diagonal matrix $\mathbf{g}_0$ of Eq. (22a), and are given by

$$g_{\lambda,0}^{(+)}(R, R') = -RR'k_\lambda j_l(k_\lambda R_<) h_l^{(+)}(k_\lambda R_>), \tag{69}$$

where we have used the Hankel function $h_l^{(+)}(x) \sim x^{-1} \exp i(x - l\pi/2)$.

The stationary state $\Psi_{pnjm_j}^{(+)}$ may be expanded as in Eq. (23), with ψ_0 replaced by the matrix state $\psi^J(R)$ which satisfies

$$[k^2 - (2m/\hbar^2)(R\mathbf{K}R^{-1} + \mathbf{H}_I + \mathbf{V}^J)]\psi^J(R) = 0. \tag{70}$$

The solution is, employing Eq. (69),

$$\psi^J(R) = \psi_0(R) + (2m/\hbar^2) \int dR' \, \mathbf{g}_0^{(+)}(R, R') \mathbf{V}^J(R') \psi^J(R'), \tag{71}$$

which gives asymptotically, for $R \to \infty$,

$$\psi_{\lambda'\lambda}^J(R) \sim -k_\lambda^{-1/2} [\delta_{\lambda'\lambda} e^{-i(k_\lambda R - l\pi/2)} - S_{\lambda'\lambda}^J e^{i(k_{\lambda'} R - l'\pi/2)}] \tag{72}$$

where $S_{\lambda'\lambda}^J$ is the scattering matrix element for the $\lambda \to \lambda'$ transition.

D. Rate of Change of Observables

Let A be an observable of the two colliding particles which does not depend explicitly on time and which is a constant of motion in the absence of interaction or, in other terms, which satisfies

$$[H_0, A] = 0. \tag{73}$$

For a nonvanishing interaction V the state of the system would be $\Psi^{(+)}(t)$ and A would change in time in accordance with

$$d\langle \Psi^{(+)}(t) | A | \Psi^{(+)}(t) \rangle / dt = (i\hbar)^{-1} \langle \Psi^{(+)}(t) | [A, V] | \Psi^{(+)}(t) \rangle. \tag{74}$$

We obtain from this the rate of change of A in a stationary state, first given by Lippmann (1965). Writing Eq. (7) in the limit when the wavepacket becomes $\Psi_\nu^{(+)} \cdot \exp(-iEt/\hbar)$,

$$
\begin{aligned}
(d\langle A\rangle/dt)_\nu &= (i\hbar)^{-1}\langle\Psi_\nu^{(+)}|[A,V]|\Psi_\nu^{(+)}\rangle \\
&= (2/\hbar)\mathrm{Im}\langle\Psi_\nu^{(+)}|AV|\Psi_\nu^{(+)}\rangle \\
&= (2/\hbar)\mathrm{Im}\langle\Psi_{0\nu}|AV|\Psi_\nu^{(+)}\rangle \\
&\quad + (2/\hbar)\mathrm{Im}\langle\Psi_\nu^{(+)}|VG_0^{(-)}AV|\Psi_\nu^{(+)}\rangle
\end{aligned}
\tag{75}
$$

where we have used Eq. (60b) for $\langle\Psi_\nu^{(+)}|$. Since A and $G_0^{(+)}$ commute, the last term may be obtained from the imaginary part of the well-known relation

$$
\lim_{\varepsilon\to 0+} G_0^{(\pm)}(E) = \mathscr{P}(E-H_0)^{-1} \mp i\pi\delta(E-H_0),
\tag{76}
$$

where $\mathscr{P}$ specifies the principal value, and one arrives at the desired result

$$
\begin{aligned}
(d\langle A\rangle/dt)_\nu &= (2/\hbar)\mathrm{Im}\langle\Psi_{0\nu}|AV|\Psi_\nu^{(+)}\rangle \\
&\quad + (2/\hbar)\langle\Psi_\nu^{(+)}|V\delta(E-H_0)AV|\Psi_\nu^{(+)}\rangle.
\end{aligned}
\tag{77}
$$

Of particular interest is the choice $A = P_{0\nu}$, in which case Eq. (74) gives the probability rate for finding the wavepacket in state $\Psi_{0\nu}$ at time t. In the limit of Eq. (76) this is the rate $R_{\nu'\nu}$ for the transition probability from $\Psi_{0\nu}$ to $\Psi_{0\nu'}$. For $\nu \neq \nu'$, Eq. (77) gives

$$
\begin{aligned}
R_{\nu'\nu} &= (2\pi/\hbar)|\langle\Psi_{0\nu'}|V|\Psi_\nu^{(+)}\rangle|^2\delta(E-E') \\
&= (2\pi/\hbar)|\langle\Psi_{0\nu'}|T^{(+)}(E)|\Psi_{0\nu}\rangle|^2\,\delta(E-E'),
\end{aligned}
\tag{78}
$$

where the limit $\varepsilon \to 0+$ is understood.

Cross sections are transition rates per unit of incoming flux. For the initial stationary state $\nu = (\mathbf{p}, n)$ the incoming flux is $(p/m)(2\pi\hbar)^{-3}$, and for a final state defined by n' and by a final momentum $\mathbf{p}'$ within the solid angle $\Delta\Omega$ one finds, with the present continuum normalization,

$$
\int_{\Delta\Omega} d\mathbf{p}'\, R_{n'n}(\mathbf{p}',\mathbf{p}) = (p/m)(2\pi\hbar)^{-3}\,\Delta\sigma_{n'n}(\mathbf{p}',\mathbf{p}).
\tag{79}
$$

Using Eq. (78) and integrating, the cross section $\Delta\sigma$ is given in the limit of infinitesimal solid angles by

$$
d\sigma_{n'n}(\mathbf{p}',\mathbf{p})/d\Omega = (2\pi)^4\hbar^2 m^2(p'/p)|\langle\Psi_{0\nu'}|T^{(+)}(E)|\Psi_{0\nu}\rangle|^2,
\tag{80}
$$

where $(p')^2/2m = E - W_{n'}$.

Other rates, such as those for momentum transfer or for population of internal states, can be similarly obtained from suitable choices of A.

III. The Partitioning Method

A. Effective Hamiltonians

Only a finite number of initial and final internal states are usually of interest in a collision process, those that are initially selected or finally detected at large R. On the other hand a complete set of internal states is required to represent $\Psi_\nu^{(+)}$. The partitioning method provides a description of the coupling between the states of interest and all others. It makes use of projection operators P which satisfy

$$P = P^\dagger, \qquad P^2 = P, \tag{81}$$

and which commute asymptotically with H,

$$\lim_{R\to\infty} [H, P] = 0. \tag{82}$$

The operator P must be defined so that it projects asymptotically only on the internal states of interest. Hence it is in general a multidimensional projector satisfying

$$P\Psi_{0\nu} \sim \Psi_{0\nu}, \tag{83}$$

but may otherwise be arbitrarily chosen.

We introduce the projector Q complementary to P, that is

$$Q = 1 - P, \tag{84}$$

which may be seen to satisfy also the relations corresponding to Eqs. (81) and (82). We next describe a partitioning method symmetric in both P and Q, defining for any operator A the decomposition

$$A = A_\mathrm{D} + A_\mathrm{N}, \tag{85a}$$

$$A_\mathrm{D} = PAP + QAQ, \tag{85b}$$

$$A_\mathrm{N} = PAQ + QAP, \tag{85c}$$

where A_D and A_N are its diagonal and nondiagonal parts. For the product of two operators A and B we find

$$(AB)_\mathrm{D} = A_\mathrm{D} B_\mathrm{D} + A_\mathrm{N} B_\mathrm{N}, \tag{86a}$$

$$(AB)_\mathrm{N} = A_\mathrm{D} B_\mathrm{N} + A_\mathrm{N} B_\mathrm{D}, \tag{86b}$$

and we notice that these relationships would remain valid if more than two projectors were introduced to partition the unity operator.

We turn to the equation defining the basic operator $G^{(+)}(E)$,

$$(E^{(+)} - H)G^{(+)}(E) = 1, \tag{87}$$

and take the diagonal and nondiagonal parts of both sides to obtain

$$(E^{(+)} - H_D)G_D^{(+)} - H_N G_N^{(+)} = 1, \tag{88a}$$

$$(E^{(+)} - H_D)G_N^{(+)} - H_N G_D^{(+)} = 0. \tag{88b}$$

From this first equation we get

$$G_N^{(+)} = (E^{(+)} - H_D)^{-1} H_N G_D^{(+)}, \tag{89}$$

and replacing in the second one

$$(E^{(+)} - \mathscr{H}_D)G_D^{(+)} = 1, \tag{90}$$

where

$$\mathscr{H}_D = H_D + H_N(E^{(+)} - H_D)^{-1} H_N \tag{91}$$

is the effective hamiltonian that determines the propagator in the subspaces of either P or Q. We also find, from Eq. (89),

$$G^{(+)}(E) = G_D^{(+)} + G_N^{(+)} = F^{(+)}(E)G_D^{(+)} \tag{92}$$

with

$$F^{(+)}(E) = 1 + (E^{(+)} - H_D)^{-1} H_N. \tag{93}$$

For stationary scattering states, the wave operator as given by Eq. (62) leads to

$$\begin{aligned} W_D^{(+)}(E) &= i\varepsilon G_D^{(+)}(E) = i\varepsilon(E^{(+)} - \mathscr{H}_D)^{-1} \\ &= 1 + G_D^{(+)}(E)(\mathscr{H}_D - E), \end{aligned} \tag{94}$$

and from Eq. (92) to

$$W^{(+)}(E) = F^{(+)}(E) W_D^{(+)}(E). \tag{95}$$

Another useful expression follows from rewriting Eq. (90) in the form

$$(E^{(+)} - H_D)G_D^{(+)} = 1 + H_N(E^{(+)} - H_D)^{-1} H_N G_D^{(+)}, \tag{96}$$

multiplying both sides by $H_N(E^{(+)} - H_D)^{-1}$, solving for $H_N G_D^{(+)}$ and replacing this in Eq. (96) to obtain

$$G_D^{(+)} = (E^{(+)} - H_D)^{-1} + (E^{(+)} - H_D)^{-1} H_N G_D^{(+)} H_N (E^{(+)} - H_D)^{-1}. \tag{97}$$

The previous partitioning method has been frequently applied in two different ways: to construct effective (or optical) potentials and to describe resonance scattering processes. In the first case one wishes to calculate cross sections for transitions within the manifold of P over a usually small range of collision energies. One then requires that P satisfy

$$P\Psi_{0\mathrm{v}} = \Psi_{0\mathrm{v}} \tag{98}$$

for all the possible free states $\Psi_{0\mathrm{v}}$, and wants to calculate

$$P\Psi_{\mathrm{v}}^{(+)} = PW^{(+)}\Psi_{0\mathrm{v}} = PW_{\mathrm{D}}^{(+)}P\Psi_{0\mathrm{v}}. \tag{99}$$

From this and from Eq. (94) it is seen that only $P\mathscr{H}_{\mathrm{D}}P = \mathscr{H}_P$ is needed for the calculation. We write it in the form

$$\begin{aligned} \mathscr{H}_P &= PHP + PHQ(E^{(+)} - QHQ)^{-1}QHP \\ &= PH_0P + \mathscr{V}_P(E) \end{aligned} \tag{100}$$

with

$$\mathscr{V}_P(E) = PVP + PHQ(E^{(+)} - QHQ)^{-1}QHP \tag{101}$$

the effective potential. In accordance with Eq. (82) it goes asymptotically to zero for large R. The inverse operator in Eq. (101) is defined only within the manifold of Q and is called a reduced Green function. The projector P is usually chosen to include not only the initial and final internal states of interest but also those with lower energy and some others with higher energy that may be strongly coupled to the first ones.

Two situations may then arise that are best discussed referring to Fig. 3. We show schematically here the spectra for the hamiltonians *PHP*, *QHQ* and *H*. The spectrum of *H* follows from those of *PHP* and *QHQ* once the coupling *PHQ* is included, and shows in general discrete levels and a continuum with embedded (quasi-discrete) levels. The hamil-

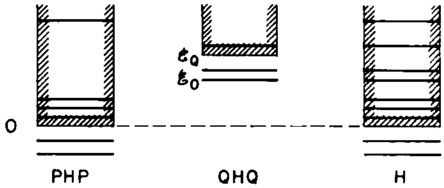


Fig. 3. Schematic spectra for the hamiltonians *PHP*, *QHQ* and *H* with a choice of *P* appropriate for effective, or optical, potentials.

tonian QHQ possesses a spectrum with a continuum starting at $\mathscr{E}_Q$. Let $\mathscr{E}_0$ be the lowest (positive) discrete eigenvalue of QHQ. We may have

$$0 < E < \mathscr{E}_0, \tag{102}$$

in which case

$$E - QHQ < 0 \tag{103}$$

and we can set $\varepsilon = 0$ in Eq. (101). Then $\mathscr{V}_P$ is real and

$$\mathscr{V}_P = PVP + PHQ(E - QHQ)^{-1}QHP < PVP, \tag{104}$$

so that coupling of the P and Q subspaces decreases the interaction within the manifold of P. As E increases the second term becomes more negative; it then changes sign between vertical asymptotes at the discrete eigenvalues of QHQ. When it reaches values

$$E \geq \mathscr{E}_Q \tag{105}$$

we must keep $\varepsilon \neq 0$. From

$$\lim_{\varepsilon \to 0+} (E^{(+)} - QHQ)^{-1} = \mathscr{P}(E - QHQ) - i\pi\, \delta(E - QHQ) \tag{106}$$

we see that $\mathscr{V}_P$ becomes a complex operator with imaginary part

$$\operatorname{Im} \mathscr{V}_P = -\pi PHQ\, \delta(E - QHQ)QHP \leq 0. \tag{107}$$

Since a complex potential with negative imaginary part leads to a negative divergence of the scattering flux (e.g., see Mott and Massey, 1965), we conclude that in this case flux is lost from the manifold of P into that of Q.

In practice $\mathscr{V}_P$ must first be found and then the scattering equations must be solved. Hence it is necessary to construct $Q(E^{(+)} - QHQ)^{-1}$, which may be a difficult task. We show in Section III, C how bounds may be given for this operator.

A second frequent application of the partitioning method is to resonance scattering which, as we shall see in Section V, involves quasi-bound states, i.e. decaying states in a time-dependent description. In this case it is convenient to first construct the projector Q. Choosing a set of orthonormalized states Φ_κ to represent the quasi-bound states at short distances and taking

$$Q = \sum_{\kappa=1}^{K} |\Phi_\kappa\rangle\langle\Phi_\kappa|, \tag{108}$$

it follows that $Q \sim 0$ for $R \to \infty$, so that $P \sim 1$ and P commutes with H in this limit. Solving the eigenvalue problem

$$(\mathscr{E}_\kappa - QHQ)|\xi_\kappa\rangle = 0, \tag{109}$$

one gets eigenstates

$$\xi_\kappa = \sum_{\kappa=1}^{K} c_{\kappa\kappa'} \Phi_{\kappa'}, \tag{110}$$

and a finite set of eigenenergies $\mathscr{E}_\kappa$. The spectra of QHQ and of PHP are shown schematically for this case in Fig. 4. The eigenvalues $\mathscr{E}_\kappa$ are positive although they correspond to bound states, because the hamiltonians QHQ and H do not have in general the same limiting spectral points, i.e. the origins of their continua are different.

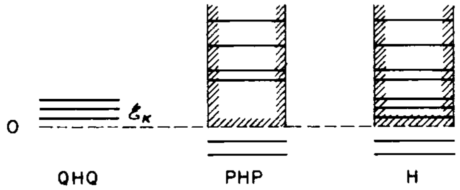


Fig. 4. Schematic spectra of the hamiltonians PHP, QHQ and H with a choice of Q appropriate for discussion of resonance scattering.

The hamiltonian $\mathscr{H}_P$ may now be written in the form

$$\mathscr{H}_P = PHP + PHQ \sum_{\kappa=1}^{K} \frac{|\xi_\kappa\rangle\langle\xi_\kappa|}{E^{(+)} - \mathscr{E}_\kappa} QHP. \tag{111}$$

It is clear that by construction the second term varies rapidly with energy E when this approaches the quasi-discrete levels, while the first term is a smooth function of E in the same region.

B. Choice of Projection Operators

Consider a collision process in which only the internal states n_i, $i = 1$ to N, are important. The projector P must asymptotically select at least these states and it must satisfy Eqs. (81), (82). The simplest choice for P is

$$P = \sum_{n_i} |u_{n_i}\rangle\langle u_{n_i}|, \tag{112}$$

in which case the effective potential $\mathscr{V}_P$ given by Eq. (101) operates in the Nth dimensional manifold of P. This operator may be written in the coordinate representation as

$$\mathscr{V}_P(\mathbf{R}, \mathbf{R}') = \sum_{n_i n_i'} |u_{n_i}\rangle \mathscr{V}_{n_i n_i'}(\mathbf{R}, \mathbf{R}') \langle u_{n_i'}|, \tag{113a}$$

$$\mathscr{V}_{n_i n_i'}(\mathbf{R}, \mathbf{R}') = V_{n_i n_i'}(\mathbf{R})\delta(\mathbf{R} - \mathbf{R}') + \sum_{nn'}{}' V_{n_i n}(\mathbf{R})[(E^{(+)} - QHQ)^{-1}]_{nn'}(\mathbf{R}, \mathbf{R}') V_{n' n_i'}(\mathbf{R}'), \tag{113b}$$

which indicates its energy-dependent and nonlocal nature. The prime in the summation means that $n(n')$ differs from $n_i(n_i')$.

Actual situations are likely to be more complex, e.g., because of symmetries, as may be illustrated with our atom–diatom example. Here we want to treat equivalently all states with the same $n_i j_i$, which requires considering simultaneously the $2j_i + 1$ states with $j_i > m_{ji} > -j_i$. In this case it is convenient to further introduce the $(ljJM)$ representation, defining for each J, M, and Π

$$P_{nj}^{JM} = \sum_{i} |\mathscr{Y}_{J\lambda}^{M}\rangle\langle\mathscr{Y}_{J\lambda}^{M}|, \tag{114}$$

where we implicitly assume that the parity of $l + j$ is that of Π. With these projectors one constructs

$$P = \sum_{JM} \sum_{n_i j_i} P_{n_i j_i}^{JM}. \tag{115}$$

The kernel of the $\mathscr{V}_P$ operator depends now only on the distances R and R' and is

$$\mathscr{V}_P(R, R') = \sum_{JM} \sum_{\lambda_i \lambda_i'} |\mathscr{Y}_{J\lambda_i}^{M}\rangle \mathscr{V}_{\lambda_i \lambda_i'}^{J}(R, R') \langle\mathscr{Y}_{J\lambda_i'}^{M}|, \tag{116a}$$

$$\begin{aligned}\mathscr{V}_{\lambda_i \lambda_i'}^{J}(R, R') = {} & V_{\lambda_i \lambda_i'}^{J}(R)\delta(R - R') \\ & + \sum_{\lambda\lambda'}{}' V_{\lambda_i \lambda}^{J}(R)[(E^{(+)} - QHQ)^{-1}]_{\lambda\lambda'}^{J}(R, R') V_{\lambda' \lambda_i'}^{J}(R'),\end{aligned} \tag{116b}$$

which could be further evaluated by means of Eq. (19c).

Returning to the general case we can also define projectors

$$P = \int_{\Delta\nu} d\nu \, |\Psi_{0\nu}\rangle\langle\Psi_{0\nu}|, \tag{117}$$

where $\Delta\nu$ would for example indicate integration over all momentum magnitudes p and sum over certain internal states n and orientation quantum numbers of $\mathbf{p}$. Choices such as these have been used (Fano, 1961; Mies, 1968) to study configuration interaction in the continuum of many-electron systems, particularly that leading to autoionization of atoms and molecules.

A similar approach may be followed to construct the projector Q for quasi-bound states. In this case we can again employ internal states, but only those that are likely to contribute to the $|\xi_\kappa\rangle$ at short distances. We can choose

$$\Phi_\kappa(\mathbf{R}, \mathbf{r}) = \phi_{kn}(\mathbf{R}) u_n(\mathbf{r}) \tag{118}$$

where ϕ_{kn} decrease exponentially for $R \to \infty$. Equation (109) becomes then a matrix eigenvalue problem. For our atom–diatom example we instead choose, to construct Q^{JM}, a set of states

$$\Phi_\kappa^{JM}(R, \Omega, \mathbf{r}) = \phi_{k\lambda}^{JM}(R)\mathscr{Y}_{J\lambda}^M(\Omega, \mathbf{r}) \tag{119}$$

with $\kappa = (k, \lambda)$. The eigenstates and eigenvalues of Eq. (109) are $\xi_\kappa^{JM}(\mathbf{R}, \mathbf{r})$ and $\mathscr{E}_\kappa{}^J$ respectively, and the second term of the hamiltonian in Eq. (111) becomes

$$\begin{aligned} \mathscr{H}_P^{(r)}(R, R') = \sum_{JM} \sum_{\lambda\lambda'}{}' \,|\mathscr{Y}_{J\lambda}^M\rangle \Big[\sum_\kappa \langle \mathscr{Y}_{J\lambda}^M | V | \xi_\kappa^{JM} \rangle_R \\ \times \langle \xi_\kappa^{JM} | V | \mathscr{Y}_{J\lambda}^M \rangle_{R'} (E^{(+)} - \mathscr{E}_\kappa^J)^{-1} \Big] \langle \mathscr{Y}_{J\lambda'}^M |, \end{aligned} \tag{120}$$

where the prime in the second summation means that the $\mathscr{Y}_{J\lambda}^M$ must be different from those in the ξ_κ^{JM}.

C. Bounds to the Effective Potentials

It is possible in certain cases to give upper and lower bounds to the effective potential defined by Eq. (101). This may be done for some choices of P and for values of E such that the reduced propagator in Eq. (101) becomes a real-valued operator. We shall consider the case where E is below the threshold $\mathscr{E}_Q$ of the continuum spectrum of QHQ. We assume that QHQ has only one discrete eigenvalue $0 < \mathscr{E}_0 < \mathscr{E}_Q$. The more general situation may be treated along similar lines.

The eigenvalue $\mathscr{E}_0$ may not be exactly known so we shall only assume that an upper bound $\mathscr{E}_0{}'$ and a lower bound $\mathscr{E}_0''$ are known, satisfying $0 < \mathscr{E}_0'' < \mathscr{E}_0{}' < \mathscr{E}_Q$. We investigate the two cases $0 < E \leq \mathscr{E}_0''$ and $\mathscr{E}_Q > \mathscr{E} \geq \mathscr{E}_0{}'$.

For $0 < E \leq \mathscr{E}_0''$ the total energy is below the whole spectrum of QHQ and we can write, with $\varepsilon = 0$,

$$0 > Q(E - \mathscr{E}_0'') > Q(E - QHQ) \tag{121}$$

from which we obtain

$$Q(E - \mathscr{E}_0'')^{-1} < Q(E - QHQ)^{-1} < 0. \tag{122}$$

Upper and lower bounds to this negative operator may be obtained following methods previously developed for both bound and scattering states. Introducing a set of normalizable states $h_i(\mathbf{R}, \mathbf{r})$, $i = 1$ to M, in the form of a row matrix

$$|\mathbf{h}\rangle = (|h_1\rangle, |h_2\rangle, \ldots, |h_M\rangle), \tag{123}$$

and requiring that $Q|\mathbf{h}\rangle = |\mathbf{h}\rangle$, one can write (Lindner and Löwdin, 1968; Miller, 1969) for any operator $A > 0$ defined in the manifold of Q, the lower bound

$$A \geq |\mathbf{h}\rangle\langle\mathbf{h}|A^{-1}|\mathbf{h}\rangle^{-1}\langle\mathbf{h}| \tag{124}$$

and for $B < 0$, the upper bound

$$B^{-1} \leq |\mathbf{h}\rangle\langle\mathbf{h}|B|\mathbf{h}\rangle^{-1}\langle\mathbf{h}|, \tag{125}$$

where we are using the notation

$$(\langle\mathbf{h}|B|\mathbf{h}\rangle)_{ij} = \langle h_i|B|h_j\rangle. \tag{126}$$

Making use of the second inequality for the operator in Eq. (122), it follows that

$$Q(E - QHQ)^{-1} \leq |\mathbf{h}\rangle\langle\mathbf{h}|E - QHQ|\mathbf{h}\rangle^{-1}\langle\mathbf{h}|, \tag{127}$$

which gives an upper bound for the reduced propagator and hence for $\mathscr{V}_P$.

A lower bound may be obtained noticing that

$$Q(\mathscr{E}_0'' - QHQ)(E - QHQ)^{-1} \geq 0, \tag{128}$$

and hence that

$$\begin{aligned} &Q(E - QHQ)^{-1} - Q(E - \mathscr{E}_0'') \\ &\qquad = -Q(E - \mathscr{E}_0'')^{-1}(\mathscr{E}_0'' - QHQ)(E - QHQ)^{-1} \geq 0. \end{aligned} \tag{129}$$

Applying Eq. (124) to the operator in Eq. (129), one obtains the lower bound

$$\begin{aligned} Q(E - QHQ)^{-1} \geq{} & Q(E - \mathscr{E}_0'')^{-1}[1 - (\mathscr{E}_0'' - QHQ)|\mathbf{k}\rangle \\ & \times \langle\mathbf{k}|(\mathscr{E}_0'' - QHQ)(E - QHQ)|\mathbf{k}\rangle^{-1}\langle\mathbf{k}|(\mathscr{E}_0'' - QHQ)] \end{aligned} \tag{130}$$

where we have introduced $|\mathbf{k}\rangle = (QHQ - \mathscr{E}_0'')^{-1}|\mathbf{h}\rangle$.

For $\mathscr{E}_Q > E \geq \mathscr{E}_0'$ the treatment is somewhat more involved, and we follow here the work of Sugar and Blankenbecler (1964). For the upper bound we make use of the approximate eigenstate $|b'\rangle = Q|b'\rangle$ which satisfies

$$\langle b'|QHQ|b'\rangle = \mathscr{E}_0'. \tag{131}$$

Noticing that now

$$\begin{aligned} &Q(E - QHQ)^{-1} - |b'\rangle\langle b'|(E - \mathscr{E}_0')^{-1} \\ &\qquad = -Q(E - \mathscr{E}_0')|b'\rangle\langle b'|(\mathscr{E}_0' - QHQ)Q(E - QHQ)^{-1} \leq 0 \end{aligned} \tag{132}$$

and using Eq. (124) on the product of the last two factors, one gets

$$\begin{aligned} Q(E - QHQ)^{-1} \le |b'\rangle\langle b'|(E - \mathscr{E}_0')^{-1}[1 - (\mathscr{E}_0' - QHQ)|\mathbf{k}\rangle \\ \times \langle\mathbf{k}|(\mathscr{E}_0' - QHQ)(E - QHQ)|\mathbf{k}\rangle^{-1} \\ \times \langle\mathbf{k}|(\mathscr{E}_0' - QHQ)]. \end{aligned} \tag{133}$$

For the lower bound, use is made of

$$J_Q' = J_Q - J_Q|b'\rangle(\mathscr{E}_0' - \mathscr{E}_Q)^{-1}\langle b'|J_Q \ge 0 \tag{134}$$

where

$$J_Q = QHQ - \mathscr{E}_Q. \tag{135}$$

From Eq. (124) one gets

$$J_Q' \ge |\mathbf{h}\rangle\langle\mathbf{h}|(J_Q')^{-1}|\mathbf{h}\rangle^{-1}\langle\mathbf{h}| = J_Q^{(l)} \tag{136}$$

and using this and Eq. (134),

$$\begin{aligned} Q(E - QHQ)^{-1} &= Q(E - \mathscr{E}_Q - J_Q)^{-1} \\ &\ge Q[E - \mathscr{E}_Q - J_Q^{(l)} - J_Q|b'\rangle(\mathscr{E}_0' - \mathscr{E}_Q)^{-1}\langle b'|J_Q]^{-1} \end{aligned} \tag{137}$$

which is the remaining bound.

Equations (127) and (130) for $0 < E \le \mathscr{E}_0''$ and Eqs. (133) and (137) for $\mathscr{E}_Q > E \ge \mathscr{E}_0'$ are the main results of this subsection. They indicate how upper and lower bounds may be calculated working only in an M-dimensional manifold. This is particularly illuminating because in order to obtain the exact effective potential one would need a complete set of states, including a continuum to describe the subspace of Q. For our atom–diatom example, continuum states would be required to account for contributions of the atom-exchange channels $AB + C$ and $AC + B$ and of the dissociation channel $A + B + C$. The present bounds would allow us to avoid using continuum states, provided the thresholds for these channels are above $\mathscr{E}_Q$.

IV. Approximation Techniques

The choices of projection operators P described in Section III, B are somewhat restrictive. They have been made on the basis of the internal states u_n, and these in turn resulted from a splitting of the total hamiltonian H suggested by its asymptotic form at large R. Other choices of P result from splitting H in ways convenient not only asymptotically but for all R. We may then write

$$H = H^{(0)} + H^{(1)}, \tag{138a}$$

$$H^{(0)} \sim K + H_{\mathrm{I}}, \tag{138b}$$

for $R \to \infty$, and we may require that

$$[H^{(0)}, P] = 0, \tag{139}$$

or equivalently $(H^{(0)})_N = 0$. Limits of validity of a choice of P and of the splitting in Eq. (138) may be found by investigating the rate of convergence of the effective hamiltonian $\mathscr{H}_{\mathrm{D}}$ in powers of $H^{(1)}$. We have now

$$\mathscr{H}_{\mathrm{D}} = H_{\mathrm{D}}^{(0)} + H_{\mathrm{D}}^{(1)} + H_{\mathrm{N}}^{(1)}(E^{(+)} - H_{\mathrm{D}}^{(0)} - H_{\mathrm{D}}^{(1)})^{-1} H_{\mathrm{N}}^{(1)}, \tag{140}$$

and rewriting Eq. (99) in the form

$$(E^{(+)} - \mathscr{H}_P) P \Psi_v^{(+)} = i\varepsilon P \Psi_{0v}, \tag{141}$$

we can obtain validity conditions by requiring that second-order terms on the left-hand side of Eq. (141) are small, or

$$PHP\Psi_v^{(+)} \gg PH^{(1)}Q(E^{(+)} - QHQ)^{-1} QH^{(1)} P \Psi_v^{(+)}, \tag{142}$$

in which case we would need only the low order terms of $\mathscr{H}_P$ to calculate physical quantities such as cross sections.

We shall consider expansions of this sort in three cases: (1) the distortion approximation, (2) the adiabatic approximation, and (3) the sudden approximation. These approximations could be used to different orders in $H^{(1)}$. In the simplest approach one would work within the subspace of P with the eigenfunctions of $PH^{(0)}P$ and would treat $PH^{(1)}P$ as a perturbation. A more accurate approach would treat $PH^{(1)}P$ exactly and the remainder of $\mathscr{H}_P$ as a perturbation. This remainder measures the error made in working only within P. The following validity conditions tell us when we can assume that this error is small. Other ways of estimating errors may be found in related work on electron scattering by atoms by Mittleman and Watson (1959) and Mittleman (1965).

A. The Distortion Approximation

In this case we choose projection operators like the ones in Section III, B,

$$P_n = |u_n\rangle\langle u_n|, \tag{143a}$$

$$P = \sum_{n=n_1}^{n_2} P_n, \tag{143b}$$

and introduce the splitting of H by defining

$$V^{(0)} = \sum_n P_n V P_n \tag{144}$$

where the sum extends here over all n, to obtain

$$H^{(0)}_{\mathrm{dis}} = K + H_{\mathrm{I}} + V^{(0)}, \tag{145a}$$

$$H^{(1)}_{\mathrm{dis}} = V - V^{(0)}. \tag{145b}$$

The hamiltonian $H^{(0)}_{\mathrm{dis}}$ accounts for distortion of relative motion due to variation of V in each state u_n, but does not couple internal states. Its eigenfunctions may readily be written in the form $\phi_{n,\mathrm{dis}}(\mathbf{R})u_n(\mathbf{r})$, where $\phi_{n,\mathrm{dis}}$ satisfies

$$[E - K_{\mathbf{R}} - W_n - V_{nn}(\mathbf{R})]\phi_{n,\mathrm{dis}}(\mathbf{R}) = 0, \tag{146}$$

and scattering or bound-state boundary conditions depending on whether $E - W_n$ is positive or negative. The effective hamiltonian $\mathscr{H}_P$ is given to second order by

$$\mathscr{H}_P = \sum_{n,n'=n_1}^{n_2} |u_n\rangle(\mathscr{H}^{(0)}_{nn'} + \mathscr{H}^{(1)}_{nn'} + \mathscr{H}^{(2)}_{nn'})\langle u_{n'}|, \tag{147a}$$

$$\mathscr{H}^{(0)}_{nn'} = \delta_{nn'}(K + W_n + V_{nn}), \tag{147b}$$

$$\mathscr{H}^{(1)}_{nn'} = V_{nn'} - \delta_{nn'} V_{nn}, \tag{147c}$$

$$\mathscr{H}^{(2)}_{nn'} = \sum_{n''}{}' \, V_{nn''} G^{(+)}_{n'',\mathrm{dis}}(E) V_{n''n'} \tag{147d}$$

where n'' in Eq. (147d) must be different from n, for $n_1 \leq n \leq n_2$. In this equation $G^{(+)}_{n,\mathrm{dis}}$ is the Green function of Eq. (146) with outgoing waves for $E - W_n \geq 0$ and exponentially decreasing behavior for $E - W_n < 0$.

To obtain a validity condition from Eq. (142), we estimate the magnitude of $G^{(+)}_{n'',\mathrm{dis}}$ as approximately $(\Delta E)^{-1}$, where ΔE is the energy change during propagation in the Q subspace. We may write for this $\Delta E \cong \Delta p \cdot v$, where v is the local relative velocity and Δp the change in relative momentum, of the order of $\hbar/(\text{interaction range}) \cong \hbar(\Delta V/\Delta R)/V$. Replacing this in the right-hand side of Eq. (142), and noticing that PHP in the left-hand side is of the order of V, we find

$$1 \gg \Delta(V^2)/[\hbar v(\Delta V/\Delta R] \tag{148}$$

where $\Delta(V^2) = PV^2P - (PVP)^2$ is the mean square deviation of V. Hence the distortion approximation is valid when coupling of the P and Q subspaces through V is small and for large velocities and large potential gradients. This approximation is appealing in that it only requires, to construct $\mathscr{H}_P$, matrix elements of V and solutions of the relatively simple Eq. (146).

The effective hamiltonian $\mathscr{H}_P$ for the atom–diatom problem may be obtained, for fixed (JM), replacing, in Eq. (147), u_n by $\mathscr{Y}_{J\lambda}^{M}$, $V_{nn'}$ by $V_{\lambda\lambda'}^{J}$, and summing over λ and λ'. Eigenfunctions of the zeroth order hamiltonian may be written as in Eq. (23) with ψ_0 replaced by a diagonal matrix χ_{dis}^{J} whose elements $\chi_{\lambda,\,\mathrm{dis}}^{J}$ are the regular solutions of

$$\left\{k^2 + \frac{d^2}{dR^2} - \frac{(l+1)l}{R^2} - \frac{2m}{\hbar^2}\left[W_{nj} + V_{\lambda\lambda}^{J}(R)\right]\right\}\chi_{\lambda,\mathrm{dis}}^{J}(R) = 0 \tag{149}$$

which satisfy, provided $k_\lambda{}^2 \geq 0$, the scattering boundary conditions

$$\chi_{\lambda,\,\mathrm{dis}}^{J} \sim k_\lambda^{-1/2}\left[e^{-i(k_\lambda R - l\pi/2)} - e^{2i\eta^{J}{}_{\lambda,\,\mathrm{dis}}}e^{i(k_\lambda R - l\pi/2)}\right] \tag{150}$$

where $\eta_{\lambda,\,\mathrm{dis}}^{J}$ is the distortion phase shift. For $k_\lambda{}^2 < 0$, bound-state boundary conditions are required and we take

$$\int dR\ |\chi_{a\lambda,\,\mathrm{dis}}^{J}(R)|^2 = N_{a\lambda,\,\mathrm{dis}}^{J}, \tag{151}$$

where a is the new quantum number of the bound triatomic system and N a normalization constant.

To complete this discussion we need the distortion Green function $G_{\mathrm{dis}}^{(+)}(\mathbf{Rr},\ \mathbf{R'r'})$ which appears in the second-order term of $\mathscr{H}_P$. We may expand this function as done in Eq. (68) to obtain radial Green functions $g_{\lambda,\mathrm{dis}}^{J}(R,\ R')$ of Eq. (149). They may be expressed in terms of $\chi_{\lambda,\mathrm{dis}}^{J}$ and of irregular solutions $\bar{\chi}_{\lambda,\mathrm{dis}}^{J}(R)$ of Eq. (149) satisfying

$$\bar{\chi}_{\lambda,\,\mathrm{dis}}^{J} \sim k_\lambda^{-1/2}e^{i(k_\lambda R - l\pi/2)} \tag{152}$$

that allow us to write

$$g_{\lambda,\mathrm{dis}}^{J}(R,\ R') = -(2i)^{-1}\chi_{\lambda,\mathrm{dis}}^{J}(R_<)\bar{\chi}_{\lambda,\mathrm{dis}}^{J}(R_>). \tag{153}$$

This expression is valid for all values of $k_\lambda{}^2$ provided we choose $\mathrm{Im}\, k_\lambda > 0$ when $k_\lambda{}^2 < 0$, in which case Eq. (152) is seen to give an exponentially decreasing function of R in the limit $R \to \infty$.

The distortion approximation for molecular collisions has been used, after its introduction by Arthurs and Dalgarno (1960), by several authors (Davidson, 1962; Roberts, 1963; Dalgarno and Henry, 1965) to describe energy transfer in simple systems such as He + H_2, H + H_2, and H_2 + H_2. Limitations of this approximation have been discussed, e.g., by Lester and Bernstein (1967).

B. The Adiabatic Approximation

This approximation applies to cases where the internal motions of colliding partners readjust smoothly at each R as if the relative motion were only a perturbation. We introduce then adiabatic states $|z_\alpha(\mathbf{R})\rangle$ and energies $\mathscr{E}_\alpha(\mathbf{R})$ which satisfy

$$[H_\mathrm{I} + V - \mathscr{E}_\alpha(\mathbf{R})]|z_\alpha(\mathbf{R})\rangle = 0, \tag{154a}$$

$$\langle z_\alpha(\mathbf{R})|z_{\alpha'}(\mathbf{R})\rangle = \delta_{\alpha\alpha'}, \tag{154b}$$

with $z_\alpha \sim u_n$, so that $\mathscr{E}_\alpha \sim W_n$, for large R. Defining projection operators

$$P_\alpha(\mathbf{R}) = |z_\alpha(\mathbf{R})\rangle\langle z_\alpha(\mathbf{R})|, \tag{155a}$$

$$P(\mathbf{R}) = \sum_{\alpha=\alpha_1}^{\alpha_2} P_\alpha(\mathbf{R}), \tag{155b}$$

we choose for the adiabatic hamiltonian splitting

$$H_\mathrm{ad}^{(0)} = \sum_\alpha P_\alpha(\mathbf{R})[K + \mathscr{E}_\alpha(\mathbf{R})]P_\alpha(\mathbf{R}), \tag{156a}$$

$$H_\mathrm{ad}^{(1)} = \sum_{\alpha\neq\alpha'} P_\alpha(\mathbf{R})KP_{\alpha'}(\mathbf{R}), \tag{156b}$$

where sums extend over all α and we have assumed that the set of projectors in Eq. (155a) is complete for each $\mathbf{R}$. Equations (156) involve matrix elements of K, which are operators of the form

$$\langle z_\alpha(\mathbf{R})|K|z_{\alpha'}(\mathbf{R})\rangle = -\frac{\hbar^2}{2m}\nabla^2\delta_{\alpha\alpha'} + C_{\alpha\alpha'}(\mathbf{R}, \nabla), \tag{157a}$$

$$C_{\alpha\alpha'}(\mathbf{R}, \nabla) = -\frac{\hbar^2}{2m}(\langle z_\alpha|\nabla^2 z_{\alpha'}\rangle + 2\langle z_\alpha|\nabla z_{\alpha'}\rangle\cdot\nabla). \tag{157b}$$

The hamiltonian in Eq. (156a) has the correct asymptotic behavior $H_\mathrm{ad}^{(0)} \sim H_0$ for $R \to \infty$. It does not couple different adiabatic states so that its eigenfunctions may readily be written in the form $\phi_{\alpha,\mathrm{ad}}(\mathbf{R})z_\alpha(\mathbf{R}, \mathbf{r})$, with $\phi_{\alpha,\mathrm{ad}}$ satisfying

$$[E + (\hbar^2/2m)\nabla^2 - \mathscr{E}_\alpha(\mathbf{R}) - C_{\alpha\alpha}(\mathbf{R}, \nabla)]\phi_{\alpha,\mathrm{ad}}(\mathbf{R}) = 0 \tag{158}$$

and scattering or bound-state boundary conditions depending on whether $E \geq \mathscr{E}_\alpha(\infty)$ or $E < \mathscr{E}_\alpha(\infty)$, respectively.

For the effective hamiltonian $\mathscr{H}_P$ we obtain, with α and α' between α_1 and α_2,

$$\mathscr{H}_P = \sum_{\alpha\alpha'} |z_\alpha(\mathbf{R})\rangle(\mathscr{H}^{(0)}_{\alpha\alpha'} + \mathscr{H}^{(1)}_{\alpha\alpha'} + \mathscr{H}^{(2)}_{\alpha\alpha'})\langle z_{\alpha'}(\mathbf{R}')|, \tag{159a}$$

$$\mathscr{H}^{(0)}_{\alpha\alpha'} = \delta_{\alpha\alpha'}\delta(\mathbf{R}-\mathbf{R}')[-(\hbar^2/2m)\nabla^2 + \mathscr{E}_\alpha(\mathbf{R}) + C_{\alpha\alpha}(\mathbf{R},\nabla)], \tag{159b}$$

$$\mathscr{H}^{(1)}_{\alpha\alpha'} = \delta(\mathbf{R}-\mathbf{R}')[C_{\alpha\alpha'}(\mathbf{R},\nabla) - \delta_{\alpha\alpha'}C_{\alpha\alpha}(\mathbf{R},\nabla)], \tag{159c}$$

$$\mathscr{H}^{(2)}_{\alpha\alpha'} = \sum_{\alpha''}{}' C_{\alpha\alpha''}(\mathbf{R},\nabla)G^{(+)}_{\alpha'',\mathrm{ad}}(\mathbf{R},\mathbf{R}')C_{\alpha''\alpha'}(\mathbf{R}',\nabla'), \tag{159d}$$

where α'' is different from α, $\alpha_1 \le \alpha \le \alpha_2$, and $G^{(+)}_{\alpha,\mathrm{ad}}$ is the Green function of Eq. (158).

We estimate the validity of the adiabatic approximation by proceeding as in Section IV, A. We put $G^{(+)}_{\alpha'',\mathrm{ad}} \cong (\Delta E)^{-1}$ with $\Delta E \cong \Delta p \cdot v$, where now $\Delta p \cong \hbar\mathscr{E}^{-1}\,\Delta\mathscr{E}/\Delta R$. Since PHP is of the order of $\mathscr{E}$ we get

$$1 \gg \Delta(C^2)/[\hbar v(\Delta\mathscr{E}/\Delta R)], \tag{160}$$

where $\Delta(C^2) = PC^2P - (PCP)^2$ with ∇ replaced by $imv/\hbar$ in Eq. (157b). Hence the approximation is valid for weak coupling of the P and Q subspaces through C and for large adiabatic energy gradients. The role of the local relative velocity is less obvious, since it appears in both numerator and denominator of Eq. (160), and must be discussed within particular applications.

The foregoing treatment requires knowledge of $|\nabla z_\alpha\rangle$ and $|\nabla^2 z_\alpha\rangle$, that may be obtained from Eqs. (154). Taking the gradient of Eqs. (154a) and (154b), we arrive at new equations and boundary conditions which may be solved for $|\nabla z_\alpha\rangle$ to get, provided the $z_\alpha(\mathbf{R}, \mathbf{r})$ have been chosen as real functions,

$$\nabla\mathscr{E}_\alpha = \langle z_\alpha|(\nabla V)|z_\alpha\rangle, \tag{161a}$$

$$|\nabla z_\alpha\rangle = Q_\alpha[\mathscr{E}_\alpha - Q_\alpha(H_\mathrm{I} + V)Q_\alpha]^{-1}Q_\alpha(\nabla V)|z_\alpha\rangle, \tag{161b}$$

where $Q_\alpha = 1 - P_\alpha$. Similarly, taking the laplacian of Eqs. (154a), (154b), one can solve for $|\nabla^2 z_\alpha\rangle$ to obtain

$$\nabla^2\mathscr{E}_\alpha = \langle z_\alpha|(\nabla^2 V)|z_\alpha\rangle + 2\langle z_\alpha|(\nabla V)|\nabla z_\alpha\rangle, \tag{162a}$$

$$\begin{aligned}|\nabla^2 z_\alpha\rangle = {}& -\langle\nabla z_\alpha|\nabla z_\alpha\rangle|z_\alpha\rangle \\ &+ Q_\alpha[\mathscr{E}_\alpha - Q_\alpha(H_\mathrm{I} + V)Q_\alpha]^{-1}Q_\alpha[(\nabla^2 V)|z_\alpha\rangle \\ &+ 2(\nabla V - \nabla\mathscr{E}_\alpha)\cdot|\nabla z_\alpha\rangle].\end{aligned} \tag{162b}$$

Equations (161) and (162) may be used to discuss further the magnitude of the coupling operators $C_{\alpha\alpha'}$. They for example show that $C_{\alpha\alpha} = \hbar^2\langle \nabla z_\alpha | \nabla z_\alpha \rangle/(2m) \geq 0$.

In order to apply the adiabatic approximation to atom–diatom problems it is necessary to separate the radial and orbital terms of the kinetic energy operator K. Only the radial term may be considered a perturbation on adiabatic states for fixed R, which must incorporate the effect of the orbital motion. We write the adiabatic states for fixed (J, M) in the form

$$z_{J\alpha}^{M}(\mathbf{R}, \mathbf{r}) = \sum_{\lambda} a_{\alpha\lambda}^{J}(R)\mathscr{Y}_{J\lambda}^{M}(\Omega, \mathbf{r}), \tag{163}$$

where the coefficients satisfy the matrix eigenvalue problem

$$\sum_{\lambda} \left\{\delta_{\lambda'\lambda}\left[W_{nj} + \frac{\hbar^2 l(l+1)}{2mR^2}\right] + V_{\lambda'\lambda}^{J}(R) - \mathscr{E}_{\alpha}{}^{J}(R)\right\} a_{\alpha\lambda}^{J}(R) = 0, \tag{164a}$$

$$\sum_{\lambda} a_{\alpha\lambda}^{J}(R)^* a_{\alpha'\lambda}^{J}(R) = \delta_{\alpha\alpha'}, \tag{164b}$$

and $a_{\alpha\lambda}^{J}(R) \sim \delta_{\alpha\lambda}$, so that $\mathscr{E}_\alpha(R) \sim W_{nj}$. These coefficients may be arranged in column matrices $\mathbf{a}_\alpha{}^J(R)$ by means of which we define the matrix projectors

$$P_{\alpha}^{JM}(R) = \mathbf{a}_\alpha{}^J(R)\mathbf{a}_\alpha{}^J(R)^\dagger \tag{165}$$

which play the role of $P_\alpha(\mathbf{R})$ in Eq. (155a). Coupling operators are now given by

$$C_{\alpha\alpha'}^{J}\left(R, \frac{d}{dR}\right) = -\frac{\hbar^2}{2m}\left[(\mathbf{a}_\alpha{}^J)^\dagger \frac{d^2\mathbf{a}_{\alpha'}^J}{dR^2} + 2(\mathbf{a}_\alpha{}^J)^\dagger \frac{d\mathbf{a}_{\alpha'}^J}{dR}\frac{d}{dR}\right]. \tag{166}$$

The zeroth order hamiltonian $H_{\text{ad}}^{(0)}$ is defined for fixed J and is such that it leaves different $\mathbf{a}_\alpha{}^J$ uncoupled. Hence its solutions may be written in the form $\chi_{\alpha,\text{ad}}^{J}(R)\mathbf{a}_\alpha{}^J(R)$ with

$$\left\{k^2 + \frac{d^2}{dR^2} - \frac{2m}{\hbar^2}\left[\mathscr{E}_\alpha(R) + C_{\alpha\alpha}\left(R, \frac{d}{dR}\right)\right]\right\}\chi_{\alpha,\text{ad}}^{J}(R) = 0. \tag{167}$$

As in Section IV, A, if $k^2 \geq \mathscr{E}_\alpha(\infty)$ one must impose scattering boundary conditions from which one obtains adiabatic phase shifts $\eta_{\alpha,\text{ad}}^{J}$, while if $k^2 < \mathscr{E}_\alpha(\infty)$ one must impose bound-state boundary conditions such as

$$\int dR\, |\chi_{\alpha\alpha,\text{ad}}^{J}(R)|^2 = N_{\alpha\alpha,\text{ad}}^{J}. \tag{168}$$

Corresponding to the radial Green functions of the distortion approximation we have now

$$g^J_{\alpha,\mathrm{ad}}(R, R') = -(2i)^{-1}\chi^J_{\alpha,\mathrm{ad}}(R_<)\bar{\chi}^J_{\alpha,\mathrm{ad}}(R_>) \tag{169}$$

where the $\bar{\chi}$ are irregular solutions of Eq. (167).

The adiabatic approximation for molecular collisions (introduced by Levine, 1968) has been applied by Muckerman and Bernstein (1970) to He + N_2. Other treatments of this approximation have been developed for atom–atom collisions by Watson (1968), Thorson (1965), and Nielsen and Dahler (1966).

C. The Sudden Approximation

This approximation is complementary to the previous one. It applies to cases where internal motions do not change appreciably during collision. The relative motion may then be treated to lowest order as if the internal coordinates $\mathbf{r}$ were fixed. We introduce eigenfunctions $\phi(\mathbf{R};\, \mathbf{r})$ of the hamiltonian $K + V(\mathbf{R}, \mathbf{r})$ and treat $\mathbf{r}$ as a parameter. For positive eigenenergies we write

$$[p^2/(2m) - K - V(\mathbf{R}, \mathbf{r})]\phi^{(+)}_{\mathbf{p}}(\mathbf{R};\, \mathbf{r}) = 0, \tag{170a}$$

$$\phi^{(+)}_{\mathbf{p}} \sim \chi_{\mathbf{p}}(\mathbf{R}) + (2\pi\hbar)^{-3/2} f(\Omega;\, \mathbf{r}) e^{ipR/\hbar}/R, \tag{170b}$$

with normalization $\langle\phi^{(+)}_{\mathbf{p}}|\phi^{(+)}_{\mathbf{p}'}\rangle = \delta(\mathbf{p} - \mathbf{p}')$ for fixed $\mathbf{r}$. For negative eigenenergies $\varepsilon_c(\mathbf{r})$ we similarly write

$$[\varepsilon_c(\mathbf{r}) - K - V(\mathbf{R}, \mathbf{r})]\phi_c(\mathbf{R};\, \mathbf{r}) = 0, \tag{171a}$$

$$\int d\mathbf{R}\, |\phi_c(\mathbf{R};\, \mathbf{r})|^2 = 1, \tag{171b}$$

and require that scattering and bound states are orthogonal. All these states form then a complete set at fixed $\mathbf{r}$.

We next introduce projection operators $P_{\mathbf{p}n}$ defined so that, for a given function $g(\mathbf{R}, \mathbf{r})$,

$$P_{\mathbf{p}n}\, g(\mathbf{R}, \mathbf{r}) = \phi^{(+)}_{\mathbf{p}}(\mathbf{R};\, \mathbf{r}) u_n(\mathbf{r}) \int d\mathbf{r}'\, u_n(\mathbf{r}')^* \int d\mathbf{R}'\, \phi^{(+)}_{\mathbf{p}}(\mathbf{R}';\, \mathbf{r}')^* g(\mathbf{R}', \mathbf{r}') \tag{172}$$

and similar projectors P_{cn} with $\phi^{(+)}_{\mathbf{p}}$ replaced by ϕ_c. With these we construct the sudden approximation hamiltonians

$$H^{(0)}_{\mathrm{sud}} = \sum_n \left[\int d\mathbf{p}\left(\frac{p^2}{2m} + W_n\right)P_{\mathbf{p}n} + \sum_c (\varepsilon_{cn} + W_n)P_{cn}\right], \tag{173a}$$

$$H^{(1)}_{\mathrm{sud}} = H - H^{(0)}_{\mathrm{sud}}, \tag{173b}$$

where

$$\varepsilon_{cn} = \int d\mathbf{r}\,|u_n(\mathbf{r})|^2 \varepsilon_c(\mathbf{r}) \tag{174}$$

is the eigenenergy of a bound state of the whole system averaged over the distribution of internal coordinates. From Eq. (173a) it is clear that the eigenfunctions of $H_{\text{sud}}^{(0)}$ are $\phi(\mathbf{R}, \mathbf{r})u_n(\mathbf{r})$, and it may be verified that $H_{\text{sud}}^{(0)} \sim K + H_{\text{I}}$, as it should be.

Making use of the completeness of the set of projectors to expand H in Eq. (173b), we could find a more explicit expression for $H_{\text{sud}}^{(1)}$. This involves matrix elements of H_{I} given by operators

$$\langle \phi_{\mathbf{p}}^{(+)} | H_{\text{I}} | \phi_{\mathbf{p}'}^{(+)} \rangle = \delta(\mathbf{p} - \mathbf{p}')H_{\text{I}} + C_{\mathbf{pp}'}(\mathbf{r}, \nabla_r), \tag{175a}$$

$$C_{\mathbf{pp}'} = -(\hbar^2/2m_{\text{I}})(\langle \phi_{\mathbf{p}}^{(+)} | \nabla_r^{\,2} \phi_{\mathbf{p}'}^{(+)} \rangle + 2\langle \phi_{\mathbf{p}}^{(+)} | \nabla_r \phi_{\mathbf{p}'}^{(+)} \rangle \cdot \nabla_r), \tag{175b}$$

with m_{I} the internal mass, and also similar expressions, $C_{c\mathbf{p}}$, C_{cc}, with one or both scattering states replaced by bound states. Defining

$$P = \sum_{n=n_1}^{n_2} \int d\mathbf{p}\, P_{\mathbf{p}n}, \tag{176}$$

we get for the effective hamiltonian $\mathscr{H}_P$ the expansion (with $n_1 \leq n$, $n' \leq n_2$)

$$\mathscr{H}_P = \sum_{n,n'} \int d\mathbf{p} \int d\mathbf{p}'\; |\phi_{\mathbf{p}}^{(+)} u_n\rangle \left(\sum_{i=0}^{2} \mathscr{H}_{\mathbf{p}n,\,\mathbf{p}'n'}^{(i)} \right) \langle \phi_{\mathbf{p}'}^{(+)} u_{n'}|, \tag{177a}$$

$$\mathscr{H}_{\mathbf{p}n,\,\mathbf{p}'n'}^{(0)} = \delta(\mathbf{p} - \mathbf{p}')\delta_{nn'}[p^2/(2m) + W_n], \tag{177b}$$

$$\mathscr{H}_{\mathbf{p}n,\,\mathbf{p}'n'}^{(1)} = \langle u_n | C_{\mathbf{pp}'} | u_{n'} \rangle = C_{nn'}(\mathbf{pp}'), \tag{177c}$$

$$\mathscr{H}_{\mathbf{p}n,\,\mathbf{p}'n'}^{(2)} = \sum_{n''}{}' \left[\int d\mathbf{p}''\, C_{nn''}(\mathbf{pp}'')G_{\mathbf{p}''n''}^{(+)} C_{n''n'}(\mathbf{p}''\mathbf{p}') + \sum_c C_{nn''}(\mathbf{p}c)G_{cn''}^{(+)} C_{n''n'}(c\mathbf{p}') \right], \tag{177d}$$

where n'' differs from n, $n_1 \leq n \leq n_2$, and we have used in Eq. (177d) the Green functions of $H_{\text{sud}}^{(0)}$ given by

$$G_{\mathbf{p}n}^{(+)}(E) = [E^{(+)} - p^2/(2m) - W_n]^{-1}, \tag{178}$$

and similarly for $G_{cn}^{(+)}$, with $p^2/(2m)$ replaced by ε_{cn}.

We estimate validity conditions as done before. The sum of zero and first-order terms has magnitude $p^2/(2m)$, and second-order terms are approximately $\Delta(C^2)/\Delta E$, with $\Delta(C^2) = PC^2P - (PCP)^2$ and $\Delta E \cong vh(\Delta V/R)/V$. Hence the condition is

$$V\Delta(C^2)/[vh(\Delta V/\Delta R)] \ll p^2/(2m) \tag{179}$$

and indicates that the sudden approximation is most appropriate for high relative velocities and when coupling of the $\phi_{\mathbf{p}}^{(+)}$ states by internal motions is small.

As mentioned in the Introduction, this approximation has been used to first order for nuclear collision problems, under the name of adiabatic (internal motion) approximation. For molecular problems, related schemes have been applied in conjunction with a semiclassical description of the relative motion (Kramer and Bernstein, 1964; Bernstein and Kramer, 1966; Fenstermaker and Bernstein, 1967; Cross, 1967, 1970).

V. Scattering Resonances

A. Resonance Energies and Widths

When two particles with internal structure collide there is transfer of energy between relative and internal motions. This may lead to temporarily excited states of the whole system that persist for times long compared with passage times through the collision region. Properties of these long-lived or quasi-bound states, such as lifetimes and conformations, depend on the mode of formation. These states may sometimes be found, on physical grounds, to originate from bound states of positive energy such as the Φ_κ introduced in Section III,A. The role of quasi-bound states in collision processes may be studied by constructing the scattering states $P\Psi^{(+)}$. These may be obtained, as shown by Eqs. (99) and (94), from $G_P^{(+)} = PG_D^{(+)}P$. Equation (97) gives

$$G_P^{(+)} = (E^{(+)} - PHP)^{-1}P + (E^{(+)} - PHP)^{-1}PHQ(E^{(+)} - \mathscr{H}_Q)^{-1} \times QHP(E^{(+)} - PHP)^{-1}, \tag{180}$$

where Q is now chosen as in Eq. (108) and

$$\mathscr{H}_Q = QHQ + QHP(E^{(+)} - PHP)^{-1}PHQ \tag{181}$$

is a non-hermitian operator, but in this case one defined only within the Kth dimensional manifold of Q. Hence $\mathscr{H}_Q$ may be represented by a matrix in a suitable basis set and we can write

$$Q(E^{(+)} - \mathscr{H}_Q)^{-1} = \mathscr{N}_Q(E^{(+)})/D(E^{(+)}) \tag{182}$$

with

$$D(E^{(+)}) = \det[Q(E^{(+)} - \mathscr{H}_Q)] \tag{183}$$

the determinant of the K by K matrix, and $\mathcal{N}_Q$ an operator in the manifold of Q, equal to the transpose of the matrix of cofactors of $Q(E^{(+)} - \mathcal{H}_Q)$ in the chosen basis. To continue it is convenient to first construct

$$\mathcal{H}_Q^{(+)}(E) = \lim_{\varepsilon \to 0+} \mathcal{H}_Q = QHQ + QHP[\mathcal{P}(E - PHP)^{-1}] \times PHQ - i\pi QHP\,\delta(E - PHP)PHQ \quad (184)$$

which is a complex, finite-dimensional operator with energy-dependent matrix elements, and later on to let E become a complex variable. In this way we assure that the imaginary part of $\mathcal{H}_Q^{(+)}$ is negative for real energies.

We introduce eigenvalues $f_\kappa(E)$ and eigenstates $|\zeta_\kappa(E)\rangle$ of $\mathcal{H}_Q^{(+)}$ which satisfy

$$[f_\kappa(E) - \mathcal{H}_Q^{(+)}(E)]\,|\zeta_\kappa(E)\rangle = 0. \quad (185)$$

From this it follows that, for real E,

$$\langle\zeta_\kappa^*(E)|\,[f_\kappa(E) - \mathcal{H}_Q^{(+)}(E)] = 0, \quad (186)$$

where the wavefunction corresponding to $|\zeta_\kappa^*\rangle$ is the complex conjugate of that for $|\zeta_\kappa\rangle$. The two last equations give

$$[f_\lambda(E) - f_\kappa(E)]\langle\zeta_\lambda^*(E)|\zeta_\kappa(E)\rangle = 0, \quad (187)$$

so that $|\zeta_\lambda^*\rangle$ and $|\zeta_\kappa\rangle$ are biorthogonal for $f_\lambda \neq f_\kappa$. This allows us to choose solutions with normalization

$$\langle\zeta_\lambda^*(E)|\zeta_\kappa(E)\rangle = \delta_{\lambda\kappa} \quad (188)$$

and, using this set as a basis, to rewrite Q as

$$Q = \sum_{\kappa=1}^{K} |\zeta_\kappa(E)\rangle\langle\zeta_\kappa^*(E)| \quad (189)$$

and to obtain

$$(E^{(+)} - \mathcal{H}_Q)Q = \sum_{\kappa=1}^{K} [E^{(+)} - f_\kappa(E)]\,|\zeta_\kappa\rangle\langle\zeta_\kappa^*|. \quad (190)$$

Now the matrix of the operator to the left is diagonal and its determinant becomes

$$D(E^{(+)}) = \prod_{\kappa=1}^{K} [E^{(+)} - f_\kappa(E)]. \quad (191)$$

It follows from Eqs. (180) and (182) that the complex roots of $D(E) = 0$ are poles of the Green function $G_P^{(+)}(E)$. We shall see in the next subsection that they determine the time behavior of resonances. These resonance poles are solutions of

$$\prod_{\kappa=1}^{K} [E_s - f_\kappa(E_s)] = 0 \rightarrow E_s = E_s' - \frac{\Gamma_s}{2} \tag{192}$$

and there may be more than K of them. To obtain the roots, we need $\mathrm{Re}\,f_\kappa$ and $\mathrm{Im}\,f_\kappa$, which we get by taking the scalar product of Eq. (185) with $\langle\zeta_\kappa|$ and then separating real and imaginary parts to arrive to

$$\mathrm{Re}\,f_\kappa(E) = \langle\zeta_\kappa|H + HP[\mathscr{P}(E - PHP)^{-1}]PH|\zeta_\kappa\rangle/\langle\zeta_\kappa|\zeta_\kappa\rangle, \tag{193a}$$

$$\mathrm{Im}\,f_\kappa(E) = -\pi\langle\zeta_\kappa|HP\,\delta(E - PHP)PH|\zeta_\kappa\rangle/\langle\zeta_\kappa|\zeta_\kappa\rangle. \tag{193b}$$

The energy E_s' and width Γ_s of the resonance s are seen to be coupled to each other and to the other resonances. If they satisfy $\Gamma_s \ll |E_s' - E'_{s\pm1}|$ we say that the s resonance is an isolated one, while if $\Gamma_s > |E_s' - E'_{s\pm1}|$ we speak of overlapping resonances. The whole range of intermediate situations may occur.

In the case of isolated resonances it is sometimes possible to identify the positive energy bound state Φ_κ to which the resonance relates. The projector Q is then one-dimensional and the resonance parameters follow from

$$E_s' = \mathrm{Re}\,f_\kappa(E_s) = f_\kappa'(E_s', \Gamma_s), \tag{194a}$$

$$\Gamma_s = -2\,\mathrm{Im}\,f_\kappa(E_s) = f_\kappa''(E_s', \Gamma_s), \tag{194b}$$

which may be solved by perturbation, variational or iterative methods. The state ζ_κ may similarly be related to Φ_κ. Furthermore, since Eqs. (194) are implicit ones, they are not restricted to weak coupling between Φ_κ and states outside the manifold of Q.

The energy dependence of collision states $P\Psi_\nu^{(+)}$, when resonances occur, may be obtained from

$$P\Psi_\nu^{(+)} = i\varepsilon G_P^{(+)}\Psi_{0\nu} = P\Psi_{\nu,\mathrm{dir}}^{(+)} + P\Psi_{\nu,\mathrm{res}}^{(+)} \tag{195}$$

where $P\Psi_{\nu,\mathrm{dir}}^{(+)}$ is a contribution from direct processes, given by the first term in Eq. (180), or

$$P\Psi_{\nu,\mathrm{dir}}^{(+)} = i\varepsilon(E^{(+)} - PHP)^{-1}\Psi_{0\nu} = [1 + (E^{(+)} - PHP)^{-1}PVP]\Psi_{0\nu}, \tag{196}$$

which is the collision state we would have found if resonances were not present. The resonance contribution $P\Psi_{\nu,\mathrm{res}}^{(+)}$ in Eq. (195) is related to the second term of Eq. (180). Making use of Eq. (190) we get

$$P\Psi_{\nu,\mathrm{res}}^{(+)} = (E^{(+)} - PHP)^{-1}PH\sum_\kappa \frac{|\zeta_\kappa\rangle\langle\zeta_\kappa{}^*|}{E^{(+)} - f_\kappa(E)}HP\Psi_{\nu,\mathrm{dir}}^{(+)}. \tag{197}$$

We could have also used the quotient in Eq. (182) to write down this term. In any case it shows that the collision probability amplitude at a point in space varies rapidly when E approaches one of the resonance energies E_s. This in turn means that cross sections vary rapidly around collision energies for which E is close to the E_s.

It is not necessary to base the treatment of resonances on the eigenfunctions of $\mathscr{H}_Q^{(+)}$. Any set of states that span the manifold of Q would be equally acceptable to construct the inverse in Eq. (182) and to obtain the roots of Eq. (192). Related treatments have been recently developed using the Fredholm determinant (Reinhardt and Szabo, 1970; Reinhardt, 1970; Heller and Reinhardt, 1972) and inner projections (Micha and Brändas, 1971; Brändas and Micha, 1972). Shapiro (1972) has discussed unimolecular breakup on the basis of the resonance poles of Green functions.

B. Time Dependence of Resonance States

Resonance processes are manifested not only through the energy dependence of cross sections but also through a time dependence peculiar to resonance states. To obtain this time dependence we make use of results in Section II, C.

Let $P\Psi_\nu^{(+)}$ be the state detected after collision, which as we have just seen is a sum of direct and resonance terms with different energy dependence. We introduce the wavepacket amplitudes c_ν and construct time-dependent states

$$P\Psi(t) = \int d\nu\, c_\nu P\Psi_\nu^{(+)} e^{-iEt/\hbar} \tag{198}$$

which will correspondingly be the sum of a direct collision wavepacket $P\Psi_{\text{dir}}(t)$ plus a resonance one $P\Psi_{\text{res}}(t)$. We wish to obtain the time dependence of this second term. To simplify the discussion we consider only an isolated resonance s. Separating the magnitude and angular variables of the momentum $\mathbf{p}$ in the set ν we write $c_\nu\, d\nu = (2\pi\hbar)^{-1} c(E, \omega)\, d\omega$, where ω stands for all variables other than E, and define $c(E, \omega)$ for all energies by taking $c(E, \omega) = 0$ for $E < 0$.

We can write

$$P\Psi(t) = P\Psi_{\text{dir}}(t) + \int_{-\infty}^{+\infty} \frac{dE}{2\pi\hbar}\, e^{-iEt/\hbar} X_\kappa(E) g_\kappa(E), \tag{199a}$$

$$X_\kappa(E) = \int d\omega\, c(E, \omega)(E^{(+)} - PHP)^{-1} PH \,|\zeta_\kappa\rangle\langle\zeta_\kappa{}^*|\, HP\Psi_{\nu,\text{dir}}^{(+)}\rangle, \tag{199b}$$

$$g_\kappa(E) = [E^{(+)} - f_\kappa(E)]^{-1}. \tag{199c}$$

Using the inverse transform of Eq. (46) this is

$$P\Psi(t) = P\Psi_{\mathrm{dir}}(t) + \int_{-\infty}^{+\infty} dt' X_\kappa(t') g_\kappa(t-t') \tag{200}$$

and we can calculate

$$g_\kappa(t) = \int_{-\infty}^{+\infty} \frac{dE}{2\pi\hbar} e^{-iEt/\hbar}[E^{(+)} - f_\kappa(E)]^{-1} \tag{201}$$

using Cauchy's residue theorem. For $t < 0$ we integrate over a contour along the real axis and the upper half-circle; for $t > 0$ along the real axis and the lower half-circle. Recalling that $\mathrm{Im} f_\kappa(E) < 0$, so that the pole E_s in Eq. (201) lies in the lower half-plane, we get

$$\begin{aligned} g_\kappa(t) &= 0, && t < 0, \\ &= i\,\mathrm{Res}\,\{e^{-iEt/\hbar}[E - f_\kappa(E)]^{-1}\}_{E=E_s}, && t > 0, \end{aligned} \tag{202}$$

and the residue follows from

$$\text{residue} = \lim_{E \to E_s} (E - E_s) e^{-iEt/\hbar}/(E - f_\kappa(E)], \tag{203}$$

expanding the denominator around E_s, which gives

$$\text{residue} = -(df_\kappa/dE)_{E_s}^{-1} \exp(-iE_s t/\hbar). \tag{204}$$

Replacing these results in Eq. (200) the wavepacket becomes

$$P\Psi(t) = P\Psi_{\mathrm{dir}}(t) - i\left(\int_{-\infty}^{t} dt'\, X_\kappa(t') e^{iE_s t'/\hbar}\right)\left(\frac{df_\kappa}{dE}\right)_{E_s}^{-1} e^{-(iE_s' + \Gamma_s/2)t/\hbar}. \tag{205}$$

Hence the resonance term decreases exponentially with time and the corresponding probability density decays with a time constant $\tau_s = \hbar/\Gamma_s$, which is the lifetime of the resonance. The time dependence of the pre-exponential factor in Eq. (205) depends on the form of the c_v coefficients. For large t we find asymptotically

$$\begin{aligned} \int_{-\infty}^{t} dt'\, X_\kappa(t') e^{iE_s t'/\hbar} &\sim \int_{-\infty}^{+\infty} \frac{dE}{2\pi\hbar} \int_{-\infty}^{+\infty} dt'\, e^{-i(E-E_s)t'/\hbar} X_\kappa(E) \\ &\sim \int_{-\infty}^{+\infty} dE\, \delta(E - E_s) X_\kappa(E) = X_\kappa(E_s), \end{aligned} \tag{206}$$

where we have used the Fourier transform representation of the δ-function. The result to the right of Eq. (206) must be obtained calculating first Eq. (199) in the limit $\varepsilon \to 0+$ and then making an analytic continuation to E_s.

C. Approximations to Resonance Parameters

In order to calculate resonance parameters, such as energies and widths of isolated resonances, it is necessary to know the solution of the collision problem. This can seldom be done exactly and we discuss here how the approximation techniques of Section IV may be applied to calculate eigenvalues and eigenstates of $\mathscr{H}_Q$. Introducing the splitting $H = H^{(0)} + H^{(1)}$ we write

$$\mathscr{H}_Q = QH^{(0)}Q + QH^{(1)}Q + \mathscr{H}_Q', \tag{207a}$$

$$\mathscr{H}_Q' = QH^{(1)}P(E^{(+)} - PH^{(0)}P - PH^{(1)}P)^{-1}PH^{(1)}Q, \tag{207b}$$

and look for approximations to $f_\kappa(E)$ and $|\zeta_\kappa(E)\rangle$ by means of perturbation theory. For simplicity we do this only for an isolated resonance s. Corresponding to Eq. (207a) we have

$$f_\kappa = f_\kappa^{(0)} + f_\kappa^{(1)} + f_\kappa', \tag{208a}$$

$$|\zeta_\kappa\rangle = |\zeta_\kappa^{(0)}\rangle + |\zeta_\kappa^{(1)}\rangle + |\zeta_\kappa'\rangle, \tag{208b}$$

where primed quantities contain terms of second and higher orders. Replacing in the eigenvalue equation we obtain to zeroth order

$$(f_\kappa^{(0)} - QH^{(0)}Q)|\zeta_\kappa^{(0)}\rangle = 0 \tag{209}$$

which we assume may be solved and gives a solution within the manifold of Q normalized to $\langle\zeta_\kappa^{(0)}|\zeta_\kappa^{(0)}\rangle = 1$. Introducing the intermediate normalization $\langle\zeta_\kappa^{(0)}|\zeta_\kappa\rangle = 1$ for the exact answer we may proceed as done in perturbation theory of bound states (e.g., see Löwdin, 1966) to obtain to second order

$$f_\kappa^{(1)} = \langle\zeta_\kappa^{(0)}|H^{(1)}|\zeta_\kappa^{(0)}\rangle, \tag{210a}$$

$$\begin{aligned} f_\kappa^{(2)}(E) = {} & \langle\zeta_\kappa^{(0)}|H^{(1)}Q\,\frac{1 - |\zeta_\kappa^{(0)}\rangle\langle\zeta_\kappa^{(0)}|}{f_\kappa^{(0)} - QH^{(0)}Q}\,QH^{(1)}|\zeta_\kappa^{(0)}\rangle \\ & + \langle\zeta_\kappa^{(0)}|H^{(1)}P(E^{(+)} - PH^{(0)}P)^{-1}PH^{(1)}|\zeta_\kappa^{(0)}\rangle, \end{aligned} \tag{210b}$$

where we have already replaced in the second term of Eq. (210b) the result for $|\zeta^{(1)}(E)\rangle$. The first term in this equation is what we would obtain to second order if we were to work only within Q. The second term is the one providing resonance parameters, which as we see are of second or higher order. Taking real and imaginary parts of f_κ we get

$$\operatorname{Re} f_\kappa(E) \cong \langle \zeta_\kappa^{(0)} | H + H^{(1)} \left(Q \frac{1 - |\zeta_\kappa^{(0)}\rangle\langle\zeta_\kappa^{(0)}|}{f_\kappa^{(0)} - QH^{(0)}Q} Q \right.$$
$$\left. + P \frac{\mathscr{P}}{E - PH^{(0)}P} P \right) H^{(1)} | \zeta_\kappa^{(0)} \rangle, \tag{211a}$$

$$\operatorname{Im} f_\kappa(E) \cong -\pi \langle \zeta_\kappa^{(0)} | H^{(1)} \delta(E - PH^{(0)}P) H^{(1)} | \zeta_\kappa^{(0)} \rangle, \tag{211b}$$

which replaced in Eqs. (194) provide equations for E_s' and Γ_s. The last term in Eq. (211a) leads to a resonance energy shift which is in many cases negligible compared with the first term. Nevertheless, for certain energies E this last term may be required for an accurate solution of the implicit Eqs. (194).

We proceed next to discuss different approximation techniques as done in Section IV. We shall not consider the sudden approximation since internal motions are likely to change during the lifetime of quasi-bound states while the sudden approximation, on the contrary, assumes that the distribution of **r** values does not change appreciably.

In the distortion approximation we use the hamiltonian splitting of Eq. (145). Defining

$$Q = \sum_{n=n_1}^{n_2} |u_n\rangle\langle u_n| \tag{212}$$

we obtain a zeroth order eigenvalue equation with eigenenergies $f_{\kappa,\mathrm{dis}}^{(0)}$, and eigenstates $\zeta_{\kappa,\mathrm{dis}}^{(0)}(\mathbf{R}, \mathbf{r}) = \phi_{\kappa,\mathrm{dis}}(\mathbf{R}) u_n(\mathbf{r})$ which satisfy

$$\left[f_{\kappa,\mathrm{dis}}^{(0)} + \frac{\hbar^2}{2m} \nabla^2 - W_n - V_{nn}(\mathbf{R}) \right] \phi_{\kappa,\mathrm{dis}}(\mathbf{R}) = 0, \tag{213a}$$

$$\int d\mathbf{R} \, |\phi_{\kappa,\mathrm{dis}}(\mathbf{R})|^2 = 1, \tag{213b}$$

where $\kappa = (k, n)$, the quantum number k numbering bound states of positive energy for fixed n, $n_1 \leq n \leq n_2$. From Eq. (145b) and Eqs. (211) it follows that

$$\operatorname{Re} f_{\kappa,\mathrm{dis}}(E) = f_{\kappa,\mathrm{dis}}^{(0)} + \sum_{n'}{}' \left[\sum_{k' \neq k} \frac{|\langle \phi_{\kappa,\mathrm{dis}} | V_{nn'} | \phi_{\kappa',\mathrm{dis}} \rangle|^2}{f_{\kappa,\mathrm{dis}}^{(0)} - f_{\kappa',\mathrm{dis}}^{(0)}} \right.$$
$$\left. + \langle \phi_{\kappa,\mathrm{dis}} | V_{nn'} \operatorname{Re} G_{n',\mathrm{dis}}^{(+)}(E) V_{n'n} | \phi_{\kappa,\mathrm{dis}} \rangle \right], \tag{214a}$$

$$\operatorname{Im} f_{\kappa,\mathrm{dis}}(E) = \sum_{n'}{}' \langle \phi_{\kappa,\mathrm{dis}} | V_{nn'} \operatorname{Im} G_{n',\mathrm{dis}}^{(+)}(E) V_{n'n} | \phi_{\kappa,\mathrm{dis}} \rangle, \tag{214b}$$

where n' differs from n, $n_1 \leq n \leq n_2$. The quotient term in Eq. (214a) is a contribution to the κth isolated resonance from other states in the Q subspace. The last term in Eq. (214a) and Eq. (214b) involve real and imaginary parts of the Green functions of Section IV, A. These depend on whether $W_{n'} \leq E$ or $W_{n'} > E$. In the first case both real and imaginary parts are different from zero, but in the second case $\operatorname{Im} G_{n'}^{(+)}(E) = 0$ which indicates that, in the distortion approximation, internal states within P that are energetically inaccessible do not contribute at any R to the resonance parameters.

When the coupling potential matrix elements $V_{nn'}$ are large it is convenient to use the adiabatic approximation. Introducing the hamiltonian splitting of Eqs. (156), this approximation to resonance parameters may be based on

$$Q = \sum_{\alpha=\alpha_1}^{\alpha_2} |z_\alpha(\mathbf{R})\rangle\langle z_\alpha(\mathbf{R})|. \tag{215}$$

Zeroth order eigenenergies $f_{\kappa,\mathrm{ad}}^{(0)}$ and eigenstates

$$\zeta_{\kappa,\mathrm{ad}}^{(0)}(\mathbf{R}, \mathbf{r}) = \phi_{\kappa,\mathrm{ad}}(\mathbf{R}) z_\alpha(\mathbf{R}, \mathbf{r})$$

satisfy now

$$\left[f_{\kappa,\mathrm{ad}}^{(0)} + \frac{\hbar^2}{2m}\nabla^2 - \mathscr{E}_\alpha(\mathbf{R}) - C_{\alpha\alpha}(\mathbf{R}, \nabla)\right]\phi_{\kappa,\mathrm{ad}}(\mathbf{R}) = 0, \tag{216a}$$

$$\int d\mathbf{R}\, |\phi_{\kappa,\mathrm{ad}}(\mathbf{R})|^2 = 1, \tag{216b}$$

where $\kappa = (k, \alpha)$, with the quantum number k ordering bound states of the whole system with positive energy and for given α, $\alpha_1 \leq \alpha \leq \alpha_2$. From Eq. (156b) and Eqs. (211) we obtain

$$\begin{aligned}\operatorname{Re} f_{\kappa,\mathrm{ad}}(E) = f_{\kappa,\mathrm{ad}}^{(0)} + \sum_{\alpha'}{}' \Bigg[\sum_{k' \neq k} \frac{|\langle\phi_{\kappa,\mathrm{ad}}|C_{\alpha\alpha'}|\phi_{\kappa',\mathrm{ad}}\rangle|^2}{f_{\kappa,\mathrm{ad}}^{(0)} - f_{\kappa',\mathrm{ad}}^{(0)}} \\ + \langle\phi_{\kappa,\mathrm{ad}}|C_{\alpha\alpha'} \operatorname{Re} G_{\alpha',\mathrm{ad}}^{(+)}(E) C_{\alpha'\alpha}|\phi_{\kappa,\mathrm{ad}}\rangle\Bigg],\end{aligned} \tag{217a}$$

$$\operatorname{Im} f_{\kappa,\mathrm{ad}}(E) = \sum_{\alpha'}{}' \langle\phi_{\kappa,\mathrm{ad}}|C_{\alpha\alpha'} \operatorname{Im} G_{\alpha',\mathrm{ad}}^{(+)}(E) C_{\alpha'\alpha}|\phi_{\kappa,\mathrm{ad}}\rangle, \tag{217b}$$

where α' differs from α, $\alpha_1 \leq \alpha \leq \alpha_2$. As before we must distinguish states α' for which $W_{n'} = \lim \mathscr{E}_{\alpha'} \leq E$ from those for which $W_{n'} > E$. In the first case both real and imaginary parts of $G_{\alpha',\mathrm{ad}}^{(+)}$ are different from zero, but in the second case $\operatorname{Im} G_{\alpha',\mathrm{ad}}^{(+)} = 0$, which indicates that now it is the adiabatic internal states $z_{\alpha'}$ within P that will not contribute to Eq. (217b) at any R whenever they are energetically inaccessible at $R \to \infty$.

Finally, resonance parameters E_s' and Γ_s may be obtained solving the pair of Eqs. (194), for example iteratively, after making use of Eqs. (211) in the distortion approximation or of Eqs. (217) in the adiabatic approximation.

D. Resonances in Atom–Diatom Collisions

We describe here results on low-lying resonance states of an atom A in a ^1S state colliding with a homonuclear diatomic B_2 in a $^1\Sigma$ state. We take as an example Xe + H_2 and discuss both the distortion and the adiabatic approximations. Similar approaches would apply to heteronuclear diatomics or polyatomics.

We consider specifically the collision of Xe with *para*-H_2 in its ground state $(nj) = (00)$ for relative kinetic energies $E_{\text{kin}} < W_{02} - W_{00}$. This is a particularly simple, and yet illustrative, case where only elastic events occur (Micha, 1967a,b). The interaction potential is of the van der Waals type, with a long-range attraction due to charge fluctuations, and a short-range repulsion. Both of these are anisotropic and contain only even Legendre polynomials of $\cos\gamma$ in their expansion [cf., Eq. (14)] because of symmetry. We take

$$V(R, r, \gamma) = V'(R, r)[1 + b'P_2(\cos\gamma)] + V''(R, r)[1 + b''P_2(\cos\gamma)], \quad (218)$$

where primes and double primes refer to short-range and long-range interactions respectively. Matrix elements of this potential in the basis $\mathscr{Y}_{j\lambda}^{M}$ are given by Eqs. (19c) and (20), which implicitly impose restrictions on angular momentum coupling so that $j' = j, j \pm 2$ in any single interaction. Hence only the $(nj) = (00)$ internal state may be detected asymptotically, since (01) is energetically accessible but cannot be excited, and we have only elastic processes. Other excited states such as (02) may nevertheless temporarily participate in the collision and lead to resonances below the excitation threshold W_{02}. Letting $E_s' < W_{02}$ be a resonance energy and $n_s j_s l_s$ quantum numbers of an intermediate quasi-bound state, we have then

$$[\text{Xe} + \text{H}_2(00)]_{E_s'l} \to [\text{Xe} + \text{H}_2(n_s j_s)]_{E_s'l_s} \to [\text{Xe} + \text{H}_2(00)]_{E_s'l}. \quad (219)$$

For the present problem we include in Q only the internal states going asymptotically to $(n_s j_s) = (02)$. All other internal states, even the dissociative ones, are included in principle in the subspace of P. But of these only the ones with even j may couple to $j_s = 2$ and in practice it is necessary to include in P only the (00) state. Proceeding with fixed

J, M and $\Pi = (-1)^J$, the number of quasi-bound states allowed by angular momentum and parity conservation is restricted so that for $J = 0$ only one with $l_s = j_s = 2$ is allowed, for $J = 1$ only two with $l_s = j_s \pm 1 = 1, 3$ are allowed, and for $J \geq 2$ only three with $|J - 2| \leq l_s \leq J + 2$ and parity $(-1)^J$ are allowed. Hence the only potential matrix elements required are those with $\lambda = (00J) = o$ and $\lambda = (02J), (02J \pm 2)$ which we indicate by c.

Turning to the distortion approximation we first define

$$Q^{JM} = \sum_c |\mathscr{Y}^M_{Jc}\rangle\langle\mathscr{Y}^M_{Jc}| \tag{220}$$

and proceeding as in the last section we write the eigenstate of Eq. (209) in the form

$$\zeta^{(0)}_{\kappa,\mathrm{dis}}(\mathbf{R}, \mathbf{r}) = \sum_{JMc} R^{-1}\chi_{s,\mathrm{dis}}(R)\mathscr{Y}^M_{Jc}(\Omega, \mathbf{r}), \tag{221}$$

where $s = (p, cJ)$, with p designating bound states of the three-atom system. The radial functions satisfy

$$\left\{f^{(0)}_{s,\mathrm{dis}} + \frac{\hbar^2}{2m}\left[\frac{d^2}{dR^2} - \frac{l_s(l_s+1)}{R^2}\right] - W_c - V^J_{cc}(R)\right\}\chi_{s,\mathrm{dis}}(R) = 0, \tag{222a}$$

$$\sum_c \int_0^\infty dR\, |\chi_{s,\mathrm{dis}}(R)|^2 = 1. \tag{222b}$$

In order to construct the real and imaginary parts of the distortion Green functions in Eq. (214) it is convenient to use real regular and irregular solutions, $\chi^{(\mathrm{r})}$ and $\bar{\chi}^{(\mathrm{r})}$, of the radial equation, so that

$$\left\{k^2 + \frac{d^2}{dR^2} - \frac{J(J+1)}{R^2} - \frac{2m}{\hbar^2}\left[W_{oo} + V^J_{oo}(R)\right]\right\}\chi^{(\mathrm{r})}_{o,\mathrm{dis}}(R) = 0 \tag{223}$$

with

$$\chi^{(\mathrm{r})}_{o,\mathrm{dis}} \sim k_o^{-1/2}[\sin(k_oR - J\pi/2) + K_{o,\mathrm{dis}}\cos(k_oR - J\pi/2)], \tag{224a}$$

$$\bar{\chi}^{(\mathrm{r})}_{o,\mathrm{dis}} \sim k_o^{-1/2}\cos(k_oR - J\pi/2), \tag{224b}$$

where $K_{o,\mathrm{dis}} = \tan \eta_{o,\mathrm{dis}}$. The previous complex radial solutions $\chi^{(+)}$ and $\bar{\chi}^{(+)}$ may be rewritten in terms of the new real ones, which gives

$$\begin{aligned}\mathrm{Re}\, g_{o,\mathrm{dis}}(R, R') = {} & \chi^{(\mathrm{r})}_{o,\mathrm{dis}}(R)K_{o,\mathrm{dis}}[1 + (K_{o,\mathrm{dis}})^2]^{-1}\chi^{(\mathrm{r})}_{o,\mathrm{dis}}(R') \\ & - \chi^{(\mathrm{r})}_{o,\mathrm{dis}}(R)\bar{\chi}^{(\mathrm{r})}_{o,\mathrm{dis}}(R'),\end{aligned} \tag{225a}$$

$$\mathrm{Im}\, g_{o,\mathrm{dis}}(R, R') = -\chi^{(\mathrm{r})}_{o,\mathrm{dis}}(R)[1 + (K_{o,\mathrm{dis}})^2]^{-1}\chi^{(\mathrm{r})}_{o,\mathrm{dis}}(R'). \tag{225b}$$

Proceeding as indicated by Eqs. (214) and (194) we can obtain resonance energies and widths. In the present problem widths and shifts are small for physical values of the anisotropies and we can replace, on the right-hand side of Eq. (194), E_s' by $f_s^{(0)}$ and Γ_s by zero. This gives

$$E'_{s,\,\mathrm{dis}} = f^{(0)}_{s,\,\mathrm{dis}} + \sum_{c'\neq c} \left| \int_0^\infty dR\, \chi_{s,\,\mathrm{dis}} V^J_{cc'} \chi_{s',\,\mathrm{dis}} \right|^2 (f^{(0)}_{s,\,\mathrm{dis}} - f^{(0)}_{s',\,\mathrm{dis}})^{-1} \tag{226a}$$

$$+ \left(\frac{2m}{\hbar^2}\right) \int_0^\infty dR \int_0^\infty dR'\, \chi_{s,\,\mathrm{dis}}(R) V^J_{co}(R) \mathrm{Re}\, g_{o,\,\mathrm{dis}}(R, R') V^J_{oc}(R') \chi_{s,\,\mathrm{dis}}(R')$$

$$\Gamma_{s,\,\mathrm{dis}} = \left(\frac{4m}{\hbar^2}\right) \left| \int_0^\infty dR\, \chi_{s,\,\mathrm{dis}} V^J_{co} \chi^{(\mathrm{r})}_{o,\,\mathrm{dis}} \right|^2 [1 + (K_{o,\,\mathrm{dis}})^2]^{-1},$$

with scattering functions calculated at $k^2 = (2m/\hbar^2) f^{(0)}_{s,\,\mathrm{dis}}$.

An adiabatic approximation may be developed along similar lines. We have in this case adiabatic internal states $|z^M_{J,\,o}(R)\rangle$ and $|z^M_{J,\,c}(R)\rangle$; we define the projector

$$Q^{JM}(\mathbf{R}) = \sum_c |z^M_{Jc}(\mathbf{R})\rangle\langle z^M_{Jc}(\mathbf{R})| \tag{227}$$

and solve the zeroth order problem for the bound states $\chi_{s,\,\mathrm{ad}}(R)$ satisfying

$$\left[f^{(0)}_{s,\,\mathrm{ad}} + \frac{\hbar^2}{2m} \frac{d^2}{dR^2} - \mathscr{E}_c{}^J(R) - C^J_{cc}\left(R, \frac{d}{dR}\right) \right] \chi_{s,\,\mathrm{ad}}(R) = 0, \tag{228a}$$

$$\sum_c \int_0^\infty dR\, |\chi_{s,\,\mathrm{ad}}(R)|^2 = 1, \tag{228b}$$

and for scattering states satisfying

$$\left\{ k^2 + \frac{d^2}{dR^2} - \frac{2m}{\hbar^2} \left[\mathscr{E}_o{}^J(R) + C^J_{oo}\left(R, \frac{d}{dR}\right) \right] \right\} \chi^{(\mathrm{r})}_{o,\,\mathrm{ad}}(R) = 0. \tag{229}$$

with boundary conditions appropriate to regular and irregular solutions, from which $K_{o,\,\mathrm{ad}} = \tan \eta_{o,\,\mathrm{ad}}$ may be extracted. The resonance parameters $E_{s,\,\mathrm{ad}}$ and $\Gamma_{s,\,\mathrm{ad}}$ are given as in Eq. (226), but in terms of adiabatic states and replacing $V^J_{ab}(R)$ by $(\mathscr{E}{}^J_a - W_a)\, \delta_{ab} + C^J_{ab}$. For example,

$$\Gamma_{s,\,\mathrm{ad}} = \left(\frac{4m}{\hbar^2}\right) \left| \int_0^\infty dR\, \chi_{s,\,\mathrm{ad}} C^J_{co} \chi^{(\mathrm{r})}_{o,\,\mathrm{ad}} \right|^2 [1 + (K_{o,\,\mathrm{ad}})^2]^{-1}, \tag{230}$$

with the scattering function evaluated at $k^2 = (2m/\hbar^2) \cdot f^{(0)}_{s,\,\mathrm{ad}}$. This approximation has been recently used by Muckerman and Bernstein (1969).

The wavefunction $P\Psi_v^{(+)}$ may be expanded in the present example as in Eq. (23), but with ψ_0 replaced now by a function $\psi_o(R)$, since only one scattering channel is involved. This function is furthermore a sum of direct and resonance contributions as follows from Eqs. (195–197),

$$\psi_o(R) = \chi_o^{(+)}(R) + \chi_{o,\,\mathrm{res}}(R), \tag{231}$$

where the second term contains three resonances, for $c = (02J)$ and $(02J \pm 2)$. Assuming that these resonances are isolated, the asymptotic form of Eq. (231) provides new phase shifts

$$\eta_o = \eta_{o,\,\mathrm{dir}} + \tan^{-1}[\tfrac{1}{2}\Gamma_s/(E - E_s')], \tag{232}$$

where the second term is a resonance phase shift, and the direct contribution $\eta_{o,\,\mathrm{dir}}$ is $\eta_{o,\,\mathrm{dis}}$ or $\eta_{o,\,\mathrm{ad}}$ depending on the approximation.

The example in this subsection is illustrated in Fig. 5, drawn for an atom–homonuclear diatomic pair and for the low energy region of the spectrum. Figure 5a shows bound states of AB_2 at negative energies for the

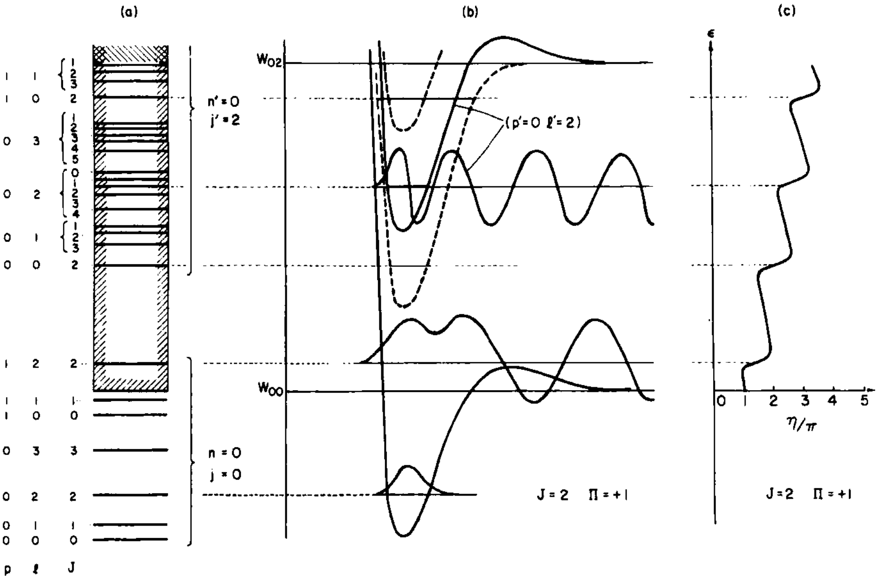


Fig. 5. (a) Spectrum of an atom–homonuclear diatomic van der Waals complex; W_{nj} is the diatomic energy and p is the quantum number for vibration of the complex. (b) From top to bottom, potentials and wave functions for the bound state, shape and compound-state resonances with $J = 2$, parity $\Pi = +1$. (c) Corresponding phase shift as a function of collision energy ε. This figure is schematic.

parent diatomic state $n = j = 0$, designated by the quantum numbers $l = J$ and p. At positive energies we find quasi-bound states corresponding to $(A, B_2{}^\dagger)$ with multiplet structures that appear for the parent diatomic state $n = 0, j = 2$ as result of angular momentum coupling. The example is one of weak coupling, and the lowest resonances are shape resonances, due to potential barrier effects. Figure 5b gives diagonal matrix elements of $V_{oo}^{0}(R) + \hbar^2 l(l+1)/(2mR^2)$ for $J = 2$ and $\Pi = +1$. The shape of the potentials depend on whether $l = J$ or $J \pm 2$, for $(A, B_2{}^\dagger)$. Wavefunctions correspond, from the top, to a compound-state resonance, a shape resonance and a bound state. In this case only elastic scattering occurs and the phase shift is shown in Fig. 5c. The three increases at high energies are subthreshold compound-state resonances that occur as E passes through a zero in the denominator of Eq. (232). The lowest increase is an orbiting resonance, which behaves in the same way.

In certain cases the atom–molecule system may show a large number of closely spaced isolated resonances or of overlapping resonances. The approach presented here may also be applied to these cases provided additional statistical approximations are made (Miller, 1970; George and Ross, 1972; Micha, 1973).

VI. Optical Potentials

A. Energy Dependence of Absorption Rates

Accurate calculation of cross sections may require a description of coupling among not only energetically accessible, or open, channels but also among these and closed ones. As a result a very large number of channels may be involved, making calculations impractical if carried out without further simplification. In some cases physical reasons indicate that only a few of the involved channels are strongly coupled while the rest play a secondary role. In these cases optical potentials may be used to advantage. They are obtained by choosing projectors P that select all channels strongly coupled to the initial one n, and by constructing from them the operators $\mathcal{V}_P$ of Section III, A.

The connection between $\mathcal{V}_P$ and transition rates may be found with the approach of Section II, D, taking the observable A equal to the operator $Q = 1 - P$, so that

$$\left(\frac{d\langle Q\rangle}{dt}\right)_\nu = (i\hbar)^{-1}\langle\Psi_\nu^{(+)}|QH - HQ|\Psi_\nu^{(+)}\rangle. \tag{233}$$

The left-hand side of Eq. (233) gives the probability rate for transition from state ν of P into any of the states of Q. We shall call this the absorption rate $R_{\nu,\mathrm{abs}}$, which is then given by

$$R_{\nu,\mathrm{abs}} = (i\hbar)^{-1}\langle\Psi_\nu^{(+)}|HP - PH|\Psi_\nu^{(+)}\rangle. \tag{234}$$

It may be transformed by means of Eq. (100), leading to

$$\begin{aligned}\mathscr{H}_P\Psi_\nu^{(+)} &= PH[P + (E^{(+)} - QHQ)^{-1}QHP]\Psi_\nu^{(+)}\\ &= PH[P + (E^{(+)} - QHQ)^{-1}Q(E^{(+)} - H)(Q - 1)]\Psi_\nu^{(+)}.\end{aligned} \tag{235}$$

Then making use of Eq. (61) we find, provided $Q\Psi_{0\nu}$ is not an eigenfunction of QHQ with eigenvalue E (a very unlikely case), that

$$\mathscr{H}_P\Psi_\nu^{(+)} = PH\Psi_\nu^{(+)}. \tag{236}$$

Replacing this expression and its adjoint in Eq. (234) one arrives at

$$R_{\nu,\mathrm{abs}} = (i\hbar)^{-1}\langle P\Psi_\nu^{(+)}|\mathscr{H}_P{}^\dagger - \mathscr{H}_P|P\Psi_\nu^{(+)}\rangle, \tag{237}$$

and hence at (Francis and Watson, 1953; Levine, 1969)

$$R_{\nu,\mathrm{abs}} = -(2/\hbar)\langle P\Psi_\nu^{(+)}|\mathrm{Im}\,\mathscr{H}_P|P\Psi_\nu^{(+)}\rangle. \tag{238}$$

This shows quite generally that the imaginary part of the effective hamiltonian measures the extent of absorption into the weakly coupled channels in Q. Following the developments in Section III, A, Im $\mathscr{H}_P$ equals Im $\mathscr{V}_P$ which is different from zero only for $E \geq \mathscr{E}_Q$ and then, as shown by Eq. (107), is a negative operator and hence consistent with the left-hand side of Eq. (238). Diagonal matrix elements of Im $\mathscr{V}_P$ with respect to internal states are negative, while off-diagonal elements may have either sign.

Equation (238) provides the basis for a semi-empirical approach. Whenever P refers asymptotically only to the internal state n, measurement of the rate $R_{\nu,\mathrm{abs}}$ gives information on the nth diagonal element of Im $\mathscr{V}_P$. This could then be used in other problems, e.g., within a distortion approximation, to incorporate absorption effects, provided the off-diagonal matrix elements of Im $\mathscr{V}_P$ are not essential.

B. Approximations to Optical Potentials

An exact determination of optical potentials would require an effort equivalent to solving the whole multichannel problem. In practice this is avoided by introducing, on physical grounds, approximations based on splittings of the total hamiltonian such as discussed in Section IV. Equation

(140) there indicates that the contribution to optical potentials from coupling of P and Q subspaces appears first to second order in $H^{(1)}$. The real and imaginary parts of the second-order term in the effective hamiltonian are

$$\mathscr{V}_P^{(r)} = \operatorname{Re} \mathscr{H}_P^{(2)} = PH^{(1)}Q[\mathscr{P}(E - QH^{(0)}Q)^{-1}]QH^{(1)}P, \tag{239a}$$

$$\mathscr{V}_P^{(i)} = \operatorname{Im} \mathscr{H}_P^{(2)} = -\pi PH^{(1)}Q\delta(E - QH^{(0)}Q)QH^{(1)}P. \tag{239b}$$

We shall work in what follows only to second order. Errors introduced by so doing could be estimated from the third-order terms in an expansion in $H^{(1)}$.

Equations (239) may readily be discussed within the distortion, adiabatic and sudden approximations. To simplify matters we consider only one-dimensional P subspaces.

In the distortion approximation we have $P = |u_n\rangle\langle u_n|$ and make use of results in Section IV, A. The collision state is $P\Psi^{(+)}_{\mathrm{v},\mathrm{dis}}$ and describes a process where the internal state remains u_n while the relative motion is affected by distortion and flux loss due to transitions from u_n. We find, for initial momentum $\mathbf{p}$,

$$P\Psi^{(+)}_{\mathrm{v},\mathrm{dis}} = \Psi^{(+)}_{\mathrm{v},\mathrm{dis}}(\mathbf{R})u_n(\mathbf{r}), \tag{240}$$

$$[p^2/(2m) - K - V_{nn} - \mathscr{V}_{n,\mathrm{dis}}]\psi^{(+)}_{\mathrm{v},\mathrm{dis}}(\mathbf{R}) = 0, \tag{241}$$

$$\mathscr{V}_{n,\mathrm{dis}}\psi^{(+)}_{\mathrm{v},\mathrm{dis}}(\mathbf{R}) = \sum_{n' \neq n} \int d\mathbf{R}'\, V_{nn'}(\mathbf{R})G^{(+)}_{n',\mathrm{dis}}(\mathbf{R}, \mathbf{R}')V_{n'n}(\mathbf{R}')\psi^{(+)}_{\mathrm{v},\mathrm{dis}}(\mathbf{R}'). \tag{242}$$

As previously discussed, the distortion Green function in Eq. (242) depends on whether $W_{n'} \geq p^2/2m + W_n$ or not. In the second case the n' channel is closed; we find $\operatorname{Im} G^{(+)}_{n',\mathrm{dis}} = 0$ and $\operatorname{Im} V_{n',\mathrm{dis}} = 0$, i.e., no contribution to absorption in accordance to Eq. (238), although there is one to distortion. When the n' channel is open we write

$$\operatorname{Im} G^{(+)}_{n',\mathrm{dis}} = -\pi\, \delta(E - K - V_{n'n'} - W_{n'}) \tag{243}$$

and express the right-hand side in terms of the complete set of states $\phi^{(-)}_{\mathrm{v},\mathrm{dis}}$, which gives

$$\begin{aligned} \operatorname{Im} G^{(+)}_{n',\mathrm{dis}}(E) &= -\pi \int d\mathbf{p}'\, \delta[E - (p')^2/(2m) - W_{n'}]\,|\phi^{(-)}_{\mathrm{v}',\mathrm{dis}}\rangle\langle\phi^{(-)}_{\mathrm{v}',\mathrm{dis}}| \\ &= -\pi m p' \int d\Omega\, |\phi^{(-)}_{\mathrm{v}',\mathrm{dis}}\rangle\langle\phi^{(-)}_{\mathrm{v}',\mathrm{dis}}|, \end{aligned} \tag{244}$$

where now $(p')^2/(2m) = E - W_{n'}$ and Ω stands for the orientation angles of $\mathbf{p}'$. Replacing Eq. (244) in the imaginary part of Eq. (242) we find the absorption rate from Eq. (238). Since we have been keeping only second-order terms it is enough to replace $\psi^{(+)}_{\nu,\mathrm{dis}}$ by $\phi^{(+)}_{\nu,\mathrm{dis}}$ in this equation, with the result

$$(R_{\nu,\mathrm{abs}})_{\mathrm{dis}} = \frac{2\pi}{\hbar} \sum_{n' \neq n} mp' \int d\Omega \, |\langle \phi^{(-)}_{\nu',\mathrm{dis}} | V_{n'n} | \phi^{(+)}_{\nu,\mathrm{dis}} \rangle|^2. \tag{245}$$

Proceeding as in Section III, D we obtain from this the differential absorption cross section

$$\left(\frac{d\sigma_{n,\mathrm{abs}}}{d\Omega}\right)_{\mathrm{dis}} = (2\pi)^4 \hbar^2 m^2 \sum_{n' \neq n} \frac{p'}{p} \, |\langle \phi^{(-)}_{\nu',\mathrm{dis}} | V_{n'n} | \phi^{(+)}_{\nu,\mathrm{dis}} \rangle|^2. \tag{246}$$

As would be expected, only open channels n' different from n contribute to this cross section.

The adiabatic approximation may be developed along similar lines, starting with $P(\mathbf{R}) = |z_\alpha(\mathbf{R})\rangle\langle z_\alpha(\mathbf{R})|$, to arrive at

$$P\Psi^{(+)}_{\nu,\mathrm{ad}} = \psi^{(+)}_{\nu,\mathrm{ad}}(\mathbf{R}) z_\alpha(\mathbf{R}, \mathbf{r}), \tag{247}$$

$$[p^2/(2m) + W_n - K - \mathscr{E}_\alpha - C_{\alpha\alpha} - \mathscr{V}_{\alpha,\mathrm{ad}}] \psi^{(+)}_{\nu,\mathrm{ad}}(\mathbf{R}) = 0, \tag{248}$$

$$\mathscr{V}_{\alpha,\mathrm{ad}} \psi^{(+)}_{\nu,\mathrm{ad}}(\mathbf{R}) = \sum_{\alpha' \neq \alpha} \int d\mathbf{R}' \, C_{\alpha\alpha'}(\mathbf{R}, \nabla) G^{(+)}_{\alpha',\mathrm{ad}}(\mathbf{R}, \mathbf{R}') C_{\alpha'\alpha}(\mathbf{R}', \nabla') \psi^{(+)}_{\nu,\mathrm{ad}}(R'), \tag{249}$$

with $\nu = (\mathbf{p}, \alpha)$. Introducing the complete set $\phi^{(-)}_{\nu.\mathrm{ad}}$ we also get

$$\mathscr{V}^{(\mathrm{i})}_{\alpha,\mathrm{ad}} = -\pi \sum_{\alpha' \neq \alpha} mp' \int d\Omega \, C_{\alpha\alpha'} | \phi^{(-)}_{\nu',\mathrm{ad}} \rangle \langle \phi^{(-)}_{\nu',\mathrm{ad}} | C_{\alpha'\alpha} \tag{250}$$

and

$$\left(\frac{d\sigma_{n,\mathrm{abs}}}{d\Omega}\right)_{\mathrm{ad}} = (2\pi)^4 \hbar^2 m^2 \sum_{\alpha' \neq \alpha} \frac{p'}{p} \, |\langle \phi^{(-)}_{\nu',\mathrm{ad}} | C_{\alpha'\alpha} | \phi^{(+)}_{\nu,\mathrm{ad}} \rangle|^2. \tag{251}$$

Finally, in the sudden approximation we make use of Eq. (172) to define

$$P = \int d\mathbf{p} \, P_{\mathbf{p}n}. \tag{252}$$

Here the optical model describes absorption from a collision state where the distribution of $\mathbf{r}$ values is always modulated by an amplitude factor $u_n(\mathbf{r})$. The eigenfunction is

$$P\Psi^{(+)}_{\mathrm{v,sud}} = \psi^{(+)}_{\mathrm{v,sud}}(\mathbf{R};\, \mathbf{r})u_n(\mathbf{r}) \tag{253}$$

and may be obtained more conveniently in the momentum representation by writing

$$\psi^{(+)}_{\mathrm{v,sud}}(\mathbf{R};\, \mathbf{r}) = \int d\mathbf{p}'\, a_n(\mathbf{p}\mathbf{p}')\phi^{(+)}_{\mathbf{p}'}(\mathbf{R};\, \mathbf{r}), \tag{254a}$$

$$\int d\mathbf{p}''\, a_n(\mathbf{p}\mathbf{p}'')^* a_n(\mathbf{p}''\mathbf{p}) = \delta(\mathbf{p} - \mathbf{p}'), \tag{254b}$$

and by making use of Eqs. (177), to get

$$[(p')^2/(2m) + W_n - E]a_n(\mathbf{p}\mathbf{p}') + \int d\mathbf{p}''\, [C_{nn}(\mathbf{p}'\mathbf{p}'') + \mathscr{V}_{n,\mathrm{sud}}(\mathbf{p}'\mathbf{p}'')]\, a_n(\mathbf{p}\mathbf{p}'') = 0, \tag{255}$$

where $\mathscr{V}_{n,\mathrm{sud}} = (\mathscr{H}^{(2)}_{nn})_{\mathrm{sud}}$. The imaginary part of this operator, required in Eq. (238), is now

$$\mathscr{V}^{(i)}_{n,\mathrm{sud}}(\mathbf{p}\mathbf{p}'') = -\pi \sum_{n' \neq n} mp' \int d\Omega\, C_{nn'}(\mathbf{p}\mathbf{p}')C_{n'n}(\mathbf{p}'\mathbf{p}''). \tag{256}$$

Approximating $\psi^{(+)}_{\mathrm{v,sud}}$ by $\phi^{(+)}_{\mathbf{p}}(\mathbf{R};\, \mathbf{r})$ in Eq. (238), or what is the same $a_n(\mathbf{p}\mathbf{p}')$ by $\delta(\mathbf{p} - \mathbf{p}')$, we get

$$\left(\frac{d\sigma_{n,\mathrm{abs}}}{d\Omega}\right)_{\mathrm{sud}} = (2\pi)^4\hbar^2 m^2 \sum_{n' \neq n} (p'/p)\,|C_{nn'}(\mathbf{p}\mathbf{p}')|^2, \tag{257}$$

where as before $\mathbf{p}'$ is within the solid angle $d\Omega$. This shows that absorption from the state $P\Psi^{(+)}_{\mathrm{v,sud}}$ in Eq. (253) occurs in the sudden approximation by an impulsive redistribution of $\mathbf{r}$ values brought about by transitions between $u_n(\mathbf{r})$ and the $u_{n'}(\mathbf{r})$ of open channels.

Other approximations to optical or effective potentials have been developed by Wolken (1972) and Rabitz (1972; Zarur and Rabitz, 1973). In the first case a complex, local, energy-dependent potential was constructed to reproduce portions of a computed multichannel scattering matrix. In the second case effective potentials were obtained that decouple orbital and rotational angular momenta, and are designed to decrease the number of coupled scattering equations. The latter effective potentials are, however, unlike the ones reviewed here in that they do not incorporate dynamical effects.

C. Applications to Atom–Diatom Systems

We take here as an example atom–diatom collisions where energy transfer occurs among translational, vibrational, and rotational motions. For an atom colliding with a diatomic in its ground state $(nj) = (00)$ at relative kinetic energies of the order of the vibrational quantum, we would find a large number of coupled vibration–rotation channels. This number would be even larger if the diatomic is in an excited (nj) state. A relatively simple situation is that where only rotational transitions $j \to j'$ occur within a given vibrational level n. We can then construct an optical potential to describe elastic and absorption scattering from the (nj) channels.

For fixed J, M, and Π we choose P in the distortion approximation to be

$$P_{nj}^{JM} = \sum_{l} |\mathcal{Y}_{Jnjl}^{M}\rangle\langle\mathcal{Y}_{Jnjl}^{M}|, \tag{258}$$

where $|J-j| \le l \le J+j$, and $(-1)^{j+l} = \Pi$. The complementary projector Q^{JM} includes all vibration–rotation channels other than the (nj) ones. How many channels must be kept in practice in Q^{JM} depends on the range of collision energies and the accuracy desired. We shall presently restrict Q^{JM} to contain only the channels (nj') with $j' \ne j$, $0 \le j' \le j_{\max}$, so that we shall be neglecting absorption into vibrational channels other than the initial one. Collision states $P\Psi_{v,\mathrm{dis}}^{(+)}$ may be expanded as shown in Eq. (23), but now $\psi_0(R)$ will be replaced by $\psi_{\mathrm{dis}}^{J}(R)$ and the summation over λ' will be restricted by $(n'j') = (nj)$. In this way one arrives at

$$[p^2/(2m) - (R\mathbf{K}R^{-1} + \mathbf{V}^J + \mathcal{V}_{\mathrm{dis}}^{J})]\psi_{\mathrm{dis}}^{J}(R) = 0, \tag{259}$$

where $p^2/2m + W_{nj} = E$, and

$$\mathcal{V}_{\lambda\lambda'}^{J}(R, R\,)_{\mathrm{dis}} = \sum_{j''l''}{}' V_{\lambda\lambda''}^{J}(R) g_{\lambda'',\mathrm{dis}}^{J}(R, R') V_{\lambda''\lambda'}^{J}(R') \tag{260}$$

with $\lambda = (njl)$, $\lambda' = (njl')$ and $j'' \ne j$. Equation (259) is a system of $N_{\mathrm{opt}} = \mathrm{Min}(2J+1,\ 2j+1)$ coupled differential equations, which must be solved, e.g., numerically, to obtain ψ_{dis}^{J}. From the asymptotic form of this function we would obtain an $N_{\mathrm{opt}} \times N_{\mathrm{opt}}$ scattering matrix $\mathbf{S}_{\mathrm{dis}}^{J}$ containing both distortion and absorption effects. On the other hand, calculation of the absorption rate $(R_{v,\mathrm{abs}})_{\mathrm{dis}}$ discussed in the previous subsection could be done from the imaginary part of Eq. (260) and would only require solutions of N_{opt} uncoupled differential equations such as those in Eq. (223), besides the solutions required, as before, to construct $\mathcal{V}_{\mathrm{dis}}^{J}$.

The adiabatic approximation to the atom–diatomic optical potentials follows along completely similar lines. One introduces the projector

$$P_{nj}^{JM}(\mathbf{R}) = \sum_{l} |z_{Jnjl}^{M}(\mathbf{R})\rangle\langle z_{Jnjl}^{M}(\mathbf{R})|, \tag{261}$$

the adiabatic matrix wave function $\psi_{\mathrm{ad}}^{J}(R)$ and Green functions $g_{\lambda'',\mathrm{ad}}^{J}$ (R, R'), and replaces the $V_{\lambda\lambda}^{J}$, (R) above by $(\mathscr{E}_{\lambda} - W_{\lambda})\,\delta_{\lambda\lambda'} + C_{\lambda\lambda'}^{J}$. From the asymptotic form of ψ_{ad}^{J} one obtains the adiabatic scattering matrix $\mathbf{S}_{\mathrm{ad}}^{J}$, after solving the same number N_{opt} of coupled differential equations.

We may compare the amount of computation presently required with that of a straightforward coupled-channel calculation. In this case the number of coupled equations is that of combinations of l and j allowed by the triangle relation for fixed J, or $N_{\mathrm{coupl}} = J(J+2)$. On the other hand the number of coupled equations in the optical model is N_{opt}, to which one must add $j_{\max}$ single equations that must be solved to construct $\mathscr{V}^{J}$. The reduction in coupled equations more than compensates for the extra work required to obtain $\mathscr{V}^{J}$ and makes optical models advantageous provided one only wishes results within the manifold of P. Furthermore, optical potentials such as that in Eq. (260) have some simplifying features. Their imaginary parts are separable in R and R', which makes them easier to handle computationally as shown, e.g., by Nakamura (1968). Other aspects of optical potentials for molecular collisions, including an eikonal approximation, have been discussed by Rotenberg (1971).

An even simpler approach consists of introducing optical potentials $V_{\mathrm{opt}}(R)$ which are both local and independent of energy for small ranges of collision velocities. In this case the absorption rate is locally given, from continuity of probability density, by a negative divergence of the scattering flux $\mathbf{J}^{(+)}$. Indicating with $\psi_{\mathbf{p}}^{(+)}(\mathbf{R})$ the collision state, Eq. (238) implies

$$\nabla \cdot \mathbf{J}^{(+)}(\mathbf{R}) = (2/\hbar)\,|\psi_{\mathbf{p}}^{(+)}(\mathbf{R})|^{2}\ \mathrm{Im}\ V_{\mathrm{opt}}(\mathbf{R}) \tag{262}$$

with Im $V_{\mathrm{opt}} \leq 0$. This approach has been followed (Marriott and Micha, 1969) for Li–HBr collisions in the relative velocity range of 10^{5} cm/sec, with HBr in the ground vibrational state. In this case vibrational excitation is forbidden by energy conservation and absorption is mostly due to rotational transitions. Calculations were performed for a potential

$$V_{\mathrm{opt}}(R) = (1 - ic)\varepsilon(R_{m}/R)^{12} - 2\varepsilon(R_{m}/R)^{6} \tag{263}$$

with $\varepsilon = 4.85 \times 10^{-14}$ erg and $R_{m} = 3.97 \times 10^{-8}$ cm. The absorption parameter c was determined by reproducing experimental total integral cross sections and was then used to predict angular distributions. Phase shifts η_l are complex in this case, so that $\eta_l = \xi_l + i\zeta_l$, and from them it is possible to calculate probabilities of absorption, or opacities, given by $p_l = 1 - \exp(4\zeta_l)$, as well as cross sections. Figure 6 shows phase shifts and opacities obtained in this way, and Fig. 7 shows differential elastic cross sections both for $c = 0$ and for $c = 0.44$. Calculations along these lines have also been carried out by Roberts and Ross (1970) and by Eu (1970), and have been reviewed by Ross and Greene (1970). It might be possible to improve upon this treatment by introducing an $N_{\text{opt}} \times N_{\text{opt}}$ matrix optical potential with local, energy-independent elements instead of Eq. (263). But this would require adjusting several optical parameters and has not yet been attempted.

A more involved application of optical potentials could be made for transitions $(nj) \rightarrow (n'j')$. In this case we would construct a projector $P^{JM} = P_{nj}^{JM} + P_{n'j'}^{JM}$ and would include all other rotational channels of

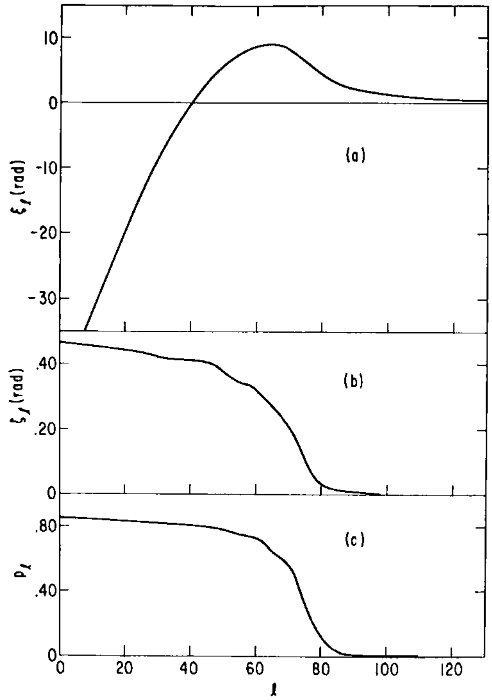


Fig. 6. (a) Real phase shifts, (b) imaginary phase shifts, and (c) opacity function at $v = 1.5 \times 10^4$ cm/sec and for $c = 0.44$ as functions of the orbital quantum number l.

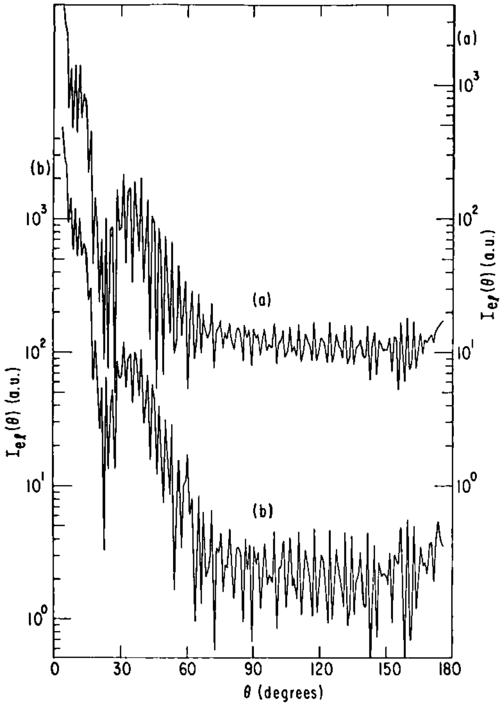


Fig. 7. (a) Differential cross section for the real potential (scale at right) and (b) for the optical potential (scale at left) versus center-of-mass scattering angle θ.

the two vibrational levels in Q^{JM}. The optical potential would then be a $2N_{\text{opt}} \times 2N_{\text{opt}}$ matrix describing elastic and absorption scattering in the (nj) and $(n'j')$ channels, transitions among them, and flux loss from them. This treatment has been partly applied to vibrational transitions in He + H_2 collisions (Micha and Rotenberg, 1970).

VII. Conclusions

Consideration of the two main subjects of this survey, resonances and optical potentials for molecular collisions, indicates that the first one has been more widely developed in the literature. This is partly due to the important role played by resonances in the description of long-lived states in crossed-beam collisions and in mechanisms of unimolecular breakup. Optical potentials have not been studied to the same extent, although they might very well become as useful in molecular processes as they have been in nuclear studies, e.g., to incorporate energy transfer effects in the treatment of direct molecular reactions.

We have treated resonance parameters here as energy-independent quantities, a feature sometimes ignored when they are introduced by means of effective hamiltonians. The bounds presented for effective potentials may provide a way to estimate inaccuracies due to use of finite sets of internal states, a problem that has so far been studied only numerically. An effort has been made to discuss the adiabatic approximation within the whole space of internal states, rather than within a finite manifold as usually done. We have developed the sudden approximation in some detail because it seems promising as a way of avoiding working with many rotational channels, a problem that has limited application of coupled-channel methods to systems of chemical interest.

As to extensions of the material presented here, the described operator approach could be easily extended, with some increase in notational complexity, to molecule–molecule collisions. A good deal remains to be done on effective hamiltonians for reactive systems. Also, observables other than transition probabilities, such as momentum transfer and population rates, could be discussed to advantage in terms of effective hamiltonians. Finally, further approximations of the semiclassical or eikonal type might bring physical insight in the form of simpler analytical results.

ACKNOWLEDGMENTS

Part of this survey was written during a visit of the author to Harvard University. He thanks Professors Alex Dalgarno, Roy Gordon, and Dudley Herschbach for their hospitality there. This work has been partly supported by a grant from the National Science Foundation and by a fellowship from the Alfred P. Sloan Foundation, for which the author is grateful.

REFERENCES

ARTHURS, A. M., and DALGARNO, A. (1960). *Proc. Roy. Soc., Ser. A***256**, 540.
AUSTERN, N., and BLAIR, J. S. (1965). *Ann. Phys. (New York)* **33**, 15.
BARDSLEY, J. N., and MANDL, F. (1968). *Rep. Progr. Phys.* **31**, 471.
BARRETT, R. (1964). *Nucl. Phys.* **51**, 27.
BATES, D. R. (1961). *In* "Quantum Theory" (D. R. Bates, ed.), Part 1, p. 287. Academic Press, New York.
BAZLEY, N. W., and FOX, D. W. (1961). *Phys. Rev.* **124**, 483.
BERNSTEIN, R. B., and KRAMER, K. H. (1966). *J. Chem. Phys.* **44**, 4473.

Brändas, E., and Micha, D. A. (1972). *J. Math. Phys.* **13**, 155.

Browne, J. C. (1971). *Advan. At. Mol. Phys.* **7**, 47.

Burke, P. G. (1965). *Advan. Phys.* **14**, 521.

Burke, P. G., Scrutton, D., Tait, J. H., and Taylor, A. J. (1969). *Proc. Phys. Soc., London* (*At. Mol. Phys.*) **2**, 1155.

Chase, D. M. (1956). *Phys. Rev.* **104**, 838.

Chen, J. C. Y. (1969). *Advan. Radiat. Chem.* **1**, 245.

Chen, J. C. Y. (1970). *Nucl. Instrum. Methods* **90**, 237.

Cross, R. J. (1967). *J. Chem. Phys.* **47**, 3724.

Cross, R. J. (1970). *In* "Molecular Beams and Reaction Kinetics" (Ch. Schlier, ed.), p. 50. Academic Press, New York.

Dalgarno, A., and Henry, R. J. W. (1965). *Proc. Phys. Soc., London* **85**, 679.

Davidson, W. D. (1962). *Discuss. Faraday Soc.* **33**, 71.

Edmonds, A. R. (1960). "Angular Momentum in Quantum Mechanics." Princeton Univ. Press, Princeton, New Jersey.

Eu, B. C. (1970). *J. Chem. Phys.* **52**, 3021.

Evans, S. A., Cohen, J. S., and Lane, N. F. (1971). *Phys. Rev. A* **4**, 2235.

Fano, U. (1961). *Phys. Rev.* **124**, 1866.

Fenstermaker, R. W., and Bernstein, R. B. (1967). *J. Chem. Phys.* **47**, 4417.

Feshbach, H. (1958a). *Ann. Phys.* (*New York*) **5**, 357.

Feshbach, H. (1958b). *Ann. Rev. Nucl. Sci.* **8**, 4.

Feshbach, H. (1962). *Ann. Phys.* (*New York*) **19**, 28.

Feshbach, H. (1967). *Ann Phys.* (*New York*) **43**, 410.

Fetter, A. L., and Watson, K. M. (1965). *Advan. Theor. Phys.* **1**, 115.

Francis, N. C., and Watson, K. M. (1953). *Phys. Rev.* **92**, 291.

Gell-Mann, M., and Goldberger, M. L. (1953). *Phys. Rev.* **91**, 398.

George, T. F., and Ross, J. (1972). *J. Chem. Phys.* **56**, 5786.

Glendenning, N. K. (1969). *In* "Nuclear Structure and Nuclear Reactions" (M. Jean and R. A. Ricci, eds.), p. 332. Academic Press, New York.

Goldberger, M. L., and Watson, K. M. (1964). "Collision Theory." Wiley, New York.

Kato, T. (1951). *Progr. Theor. Phys.* **6**, 394.

Kolos, W. (1970). *Advan. Quantum Chem.* **5**, 99.

Kramer, K. H., and Bernstein, R. B. (1964). *J. Chem. Phys.* **40**, 200.

Heller, E. J., and Reinhardt, W. P. (1972). *Phys. Rev. A* **5**, 757.

Hodgson, P. E. (1971). "Nuclear Reactions and Nuclear Structure." Oxford Univ. Press (Clarendon), London and New York.

Lane, A. M., and Robson, D. (1966). *Phys. Rev.* **151**, 774.

Lester, W. A., and Bernstein, R. B. (1967). *Chem. Phys. Lett.* **1**, 207 (erratum, p. 347).

Levine, R. D. (1968). *J. Chem. Phys.* **49**, 51.

Levine, R. D. (1969). "Quantum Mechanics of Molecular Rate Processes." Oxford Univ. Press, London and New York.

Levine, R. D. (1970). *Accounts Chem. Res.* **3**, 273.

Levine, R. D., Johnson, B. R., Muckerman, J. T., and Bernstein, R. B. (1968). *J. Chem. Phys.* **49**, 56.

Levine, R. D., Johnson, B. R., and Bernstein, R. B. (1969). *J. Chem. Phys.* **50**, 1694.

Lindner, P., and Löwdin, P. O. (1968). *Int. J. Quantum Chem.* **2S**, 161.

Lippmann, B. A. (1965). *Phys. Rev. Lett.* **15**, 11.

Lippmann, B. A., and Schwinger, J. (1950). *Phys. Rev.* **79**, 469.

Löwdin, P. O. (1963). *J. Mol. Spectrosc.* **10**, 12.

Löwdin, P. O. (1966). *In* "Perturbation Theory and its Applications in Quantum Mechanics" (C. H. Wilcox, ed.), p. 255. Wiley, New York.

Löwdin, P. O. (1967). *Advan. Quantum Chem.* **3**, 323.

Marriott, R., and Micha, D. A. (1969). *Phys. Rev.* **180**, 120.

Messiah, A. (1961). "Quantum Mechanics," Vol. 1. Wiley (Interscience), New York.

Micha, D. A. (1967a). *Chem. Phys. Lett.* **1**, 13.

Micha, D. A. (1967b). *Phys. Rev.* **162**, 88.

Micha, D. A. (1969). *J. Chem. Phys.* **50**, 722.

Micha, D. A. (1973). *Accounts Chem. Res.* **6**, 138.

Micha, D. A., and Brändas, E. (1971). *J. Chem. Phys.* **55**, 4793.

Micha, D. A., and Rotenberg, M. (1970). *Chem. Phys. Lett.* **6**, 79 [erratum, *Chem. Phys. Lett.* **11**, 626 (1971)].

Mies, F. H. (1968). *Phys. Rev.* **175**, 164.

Miller, W. H. (1969). *J. Chem. Phys.* **50**, 2758.

Miller, W. H. (1970). *J. Chem. Phys.* **52**, 5.

Mittlemann, M. H. (1965). *Advan. Theor. Phys.* **1**, 283.

Mittleman, M. H., and Watson, K. M. (1959). *Phys. Rev.* **116**, 920.

Mott, N. F., and Massey, H. S. W. (1965). "Theory of Atomic Collisions," 3rd Ed. Oxford Univ. Press, London and New York.

Muckerman, J. T., and Bernstein, R. B. (1970). *J. Chem. Phys.* **52**, 606.

Nakamura, H. (1968). *J. Phys. Soc. (Jap.)* **24**, 1353.

Newton, R. G. (1966). "Scattering Theory of Waves and Particles." McGraw-Hill, New York.

Newton, R. G., and Fonda, L. (1960). *Ann. Phys. (New York)* **9**, 416.

Nielsen, S. E., and Dahler, J. S. (1966). *J. Chem. Phys.* **45**, 4060.

Nikitin, E. E. (1968). *In* "Chemische Elementarprozesse" (H. Hartmann, ed.), p. 43. Springer-Verlag, Berlin and New York.

O'Malley, T. F. (1971). *Advan. At. Mol. Phys.* **7**, 223.

Pack, R. T., and Hirschfelder, J. O. (1968). *J. Chem. Phys.* **49**, 4009.

Pack, R. T., and Hirschfelder, J. O. (1970). *J. Chem. Phys.* **52**, 521.

Rabitz, H. (1972). *J. Chem. Phys.* **57**, 1718.

Reinhardt, W. P. (1970). *Phys. Rev. A* **2**, 1767.

Reinhardt, W. P., and Szabo, A. (1970). *Phys. Rev. A* **1**, 1162.

Riesenfeld, W. B., and Watson, K. M. (1956). *Phys. Rev.* **104**, 492.

Roberts, C. S. (1963). *Phys. Rev.* **131**, 209.

Roberts, R. E., and Ross, J. (1970). *J. Chem. Phys.* **52**, 1464.

Robson, D., and Lane, A. M. (1967). *Phys. Rev.* **161**, 982.

Roman, P. (1965). "Advanced Quantum Theory." Addison-Wesley, Reading, Massachusetts.

Ross, J., and Greene, E. F. (1970). *In* "Molecular Beams and Reaction Kinetics" (C. Schlier, ed.), p. 86. Academic Press, New York.

Rotenberg, M. (1971). *Phys. Rev. A* **4**, 220.

Shapiro, M. (1972). *J. Chem. Phys.* **56** 2582.

Smith, F. T. (1969). *Phys. Rev.* **179**, 111.

Smith, K. (1966). *Rep. Progr. Phys.* **29**, 373.

SPRUCH, L. (1969). *In* "Atomic Collision Processes" (S. Geltman, K. T. Mahanthappa, and W. E. Brittin, eds.), Lectures in Theoretical Physics, Vol. XI-C, p. 77. Gordon & Breach, New York.

SUGAR, R., and BLANKENBECLER, R. (1964). *Phys. Rev. B* **136**, 472.

TAKANAYAGI, K. (1965). *Advan. At. Mol. Phys.* **1**, 149.

TAYLOR, H. (1970). *Advan. Chem. Phys.* **18**, 91.

THORSON, W. R. (1965). *J. Chem. Phys.* **42**, 3878.

VON SEGGERN, M., and TOENNIES, J. P. (1969). *Z. Phys.* **211**, 35.

WATSON, K. M. (1957). *Phys. Rev.* **105**, 1388.

WATSON, K. M. (1968). *In* "Properties of Matter under Unusual Conditions" (H. Mark, ed.), p. 327. Wiley, New York.

WEINHOLD, F. (1972). *Advan. Quantum Chem.* **6**, 299.

WOLKEN, C., JR. (1972). *J. Chem. Phys.* **56**, 2591.

ZARUR, G., and RABITZ, H. (1973). *J. Chem. Phys.* **59**, 943.

Subject Index

A 4
B 5
C 6
D 7
E 8
F 9
G 0
H 1
I 2
J 3