ADVANCES IN QUANTUM CHEMISTRY

Volume 24

Per-Olov Löwdin

ADVANCES IN
QUANTUM CHEMISTRY

EDITOR-IN-CHIEF

PER-OLOV LÖWDIN

PROFESSOR EMERITUS

QUANTUM CHEMISTRY GROUP
UPPSALA UNIVERSITY
UPPSALA, SWEDEN

AND

QUANTUM THEORY PROJECT
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA

ASSOCIATE EDITORS

JOHN R. SABIN AND MICHAEL C. ZERNER

QUANTUM THEORY PROJECT
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA

VOLUME 24

ACADEMIC PRESS, INC.

Harcourt Brace Jovanovich, Publishers

San Diego New York Boston London Sydney Tokyo Toronto

Academic Press Rapid Manuscript Reproduction

Academic Press, Inc.
1250 Sixth Avenue, San Diego, California 92101-4311

United Kingdom Edition published by
ACADEMIC PRESS LIMITED
24-28 Oval Road, London NW1 7DX

Library of Congress Catalog Number: 64-8029

International Standard Book Number: 0-12-034824-1

PRINTED IN THE UNITED STATES OF AMERICA
92 93 94 95 96 97 BC 9 8 7 6 5 4 3 2 1

Contributors

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

Emili Besalú (117), Institute of Computational Chemistry, Faculty of Sciences, University of Girona, 17071 Girona, Spain

Ramon Carbó (117), Institute of Computational Chemistry, Faculty of Sciences, University of Girona, 17071 Girona, Spain

Sherif El-Basil (239), Faculty of Pharmacy, Cairo 11562, Egypt

Per-Olov Löwdin (79), Uppsala Quantum Chemistry Group, Uppsala University, 75120 Uppsala, Sweden; and Quantum Theory Project, University of Florida, Gainesville, Florida 32611

István Mayer (79), Central Research Institute for Chemistry, Hungarian Academy of Sciences, H-1525 Budapest, Hungary

Milan Randić (239), Department of Mathematics and Computer Science, Drake University, Des Moines, Iowa 50311

Yves G. Smeyers (5), Instituto de Estructura de la Materia (CSIC), E-28006, Spain

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, and today the new field can boast of both its conceptual framework and 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 and the referees has been so encouraging that a series of new volumes is being prepared. However, in order to control production costs and speed publication time, a new format involving camera-ready manuscripts is being used from Volume 20. A special announcement about the new format was enclosed in that volume (page xiii).

In the volumes to come, special attention will be devoted to the following subjects: the quantum theory of closed states, particularly the electronic structure of atoms, molecules, and crystals; the quantum theory of scattering states, dealing also with the theory of chemical reactions; the quantum theory of time-dependent phenomena, including the problem of electron transfer and radiation theory; molecular dynamics; statistical mechanics and general quantum statistics; condensed matter theory in general; quantum biochemistry and quantum pharmacology; the theory of numerical analysis and computational techniques.

As to the content of Volume 24, the Editors would like to thank the authors for their contributions, which give an interesting picture of part of the current state of the art of the quantum theory of matter: From an introduction to the group theory of non-rigid molecules, over a study of the general Hartree–Fock method and a treatment of the calculation of many-center integrals using Cartesian exponential-type orbitals, to an exploration of the mathematical concept of equivalence between objects of importance in chemistry and physics, including chemical graph theory.

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

INTRODUCTION TO GROUP THEORY FOR NON-RIGID MOLECULES

Yves G. Smeyers
Instituto de Estructura de la Materia (CSIC),
C/Serrano 119, E-28006, Madrid ,Spain

Octubre 15, 1991

ADVANCES IN QUANTUM CHEMISTRY
VOLUME 24

LIST OF CONTENTS

1.INTRODUCTION

A molecule may be regarded as an ensemble of atoms in a stable configuration of minimal energy in the space of the nuclear configurations. When this ensemble contains an asymmetric atom, the molecule possesses two enantiomorphic isoenergetic configurations corresponding to two optical isomers. In addition, when the ensemble contains identical but *discernible* nuclei, the molecule possesses other isoenergetic configurations, which correspond to different isomeric forms, such as rotational, positional isomers.

Whenever the different configurations are separated by relatively low energy barriers, large amplitude internal conversion movements, which transform one configuration into another isoenergetic one, can occur. These intramolecular movements can be internal rotations, pseudo-rotations and inversions or still more complex motions.

A molecule undergoing such large amplitude movements, between various possible configurations, is known as a **non-rigid molecule**. Because of this *deformability*, the non-rigid molecules exhibit some interesting properties of intramolecular dynamics, which can be studied more easily resorting to Group Theory.

For many years spectroscopists have been using the symmetry point groups for studying vibration and single torsion modes in some non-rigid molecules.

Some decades ago, Swalen and Costain(1959), Myers and Wilson (1960) have extended the use of the symmetry point groups for studying the double internal rotation in acetone [1-2]. Dreizler generalized their considerations to two C_{3v} rotor molecules with frames of a lower symmetry than C_{2v}, and deduced their character tables [3].

Simultaneously, Hougen [4] showed that the overall rotation energy levels of a rigid molecule might be classified according to the irreducible representations of the symmetry point group of the molecule. On the other hand, Longuet-Higgins [5] pointed out that the Hougen's theory was not sufficient to classify the energy levels of a non-rigid molecule. In order to study these deformable systems, Longuet-Higgins proposed the use of a more general group based on permutations and permutations-inversions of identical nuclei, he called the **Molecular Symmetry Group** [5].

Stone applied the theory of Longuet-Higgins to deduce the character tables for the multiple internal rotation in neopentane and in octahedral hexaammonium metallic complexes [6]. Dalton examined the use of the permutation-inversion groups for determining statistical weights and selection rules for radiative processes in non-rigid systems [7]. Many applications of the Molecular Symmetry Groups have been reviewed later by Bunker [8,9].

The abstract character of the Molecular Symmetry Group of Longuet- Higgins does not make easy the physical interpretation of the interconversion

operations. Therefore, Altmann [10-11] proposed an alternative theory which retains in some way the Schönfliess' formalism to describe the intramolecular conversion motions. In this theory, the general group, called Schrödinger Supergroup, is expressed as a semi-direct product of two subgroups. The first subgroup is the symmetry point group of the molecule. The second one is the set of the operations which correspond to the intramolecular conversion movements.

Beside spectroscopists, organic chemists have successfully introduced the symmetry point groups in chemical reactivity [12,13]. The use of the non-rigid group in this field, however, has scarcely been explored [14-16]. In addition to spectroscopy and organic chemistry, the non-rigid groups were also introduced in crystallography [17], and more recently non-rigid molecules embedded in a crystal lattice were considered [18].

It was sometimes believed, in the scientific literature, that the Molecular Symmetry Group of Longuet-Higgins and the Schrödinger Supergroup of Altmann were isomorphic, i.e., both theories were equivalent [19]. We shall see, however, that is not generally true, especially when some symmetry is retained in the molecule.

Woodman investigated the possibility of formulating the Molecular Symmetry Groups as products of subgroups [20]. For the most general case, no such formulation is possible. When the non-rigidity proceeds from internal rotations about one or more axes with respect to a reference frame, the Molecular Symmetry Group may be written as a semidirect product of an invariant torsional subgroup, and a frame subgroup. In such a case, the Molecular Symmetry Group should be isomorphic to the Schrödinger Supergroup.

In this paper, both theories will be briefly reviewed presenting their differences. From these comparisons, the **Non-rigid Molecule Group** (NRG) will be stricktly defined as the complete set of the molecular conversion operations which commute with a given Hamiltonian operator [21]. The operations of such a set may be written either in terms of permutations and permutations-inversions, just as in the Longuet-Higgins formalism, or either in terms of physical operations just as in the formalism of Altmann. But, the order, the structure, the symmetry properties of the group will depend exclusively on the Hamiltonian operator considered.

In the **NRG** theory developed here, however, the *formalism* of Altmann will be adopted because of its simple physical interpretation and also because of its large flexibility and generality. For example, the concept of *local groups* for an approximate treatment of non-rigid systems can be more easily introduced into the Altmann's formalism, when the Longuet-Higgins'one is completely inadapted [21,22].

In order to introduce this formalism, very simple applications of the Non-

Rigid Molecule Group theory will be given in the main part of this paper. For example, symmetry adapted potential energy function for internal molecular large amplitude motions will be deduced. Symmetry eigenvectors which factorize the Hamiltonian matrix in boxes will be derived. In the last section, applications to problems of physical interest will be forwarded. For example, conformational dependencies of molecular parameters as a function of temperature will be determined. Selection rules, as well as, torsional far infrared spectrum band structure calculations will be predicted. Finally, the torsional band structures of electronic spectra of flexible molecules will be presented.

2.THE NON–RIGID MOLECULE GROUP THEORIES

2.1. General Theory

Before commencing our study of the Group Theory for Non-Rigid Molecules, it is convenient to outline that the Group Theory for Non-Rigid Molecules is defined in the framework of the Born-Oppenheimer approximation, in which the movements of the electrons and nuclei are considered separately [23]. In addition, in this theory, the identical nuclei are regarded as distinguishable and labelled entities, although this discernability will disappear in the numerical results.

It is well known, that the full electrostatic Hamiltonian operator may be written, in terms of an electronic Hamiltonian operator and a nuclear kinetic operator:

$$H(X,x) = H_e(X,x) + T(X) \tag{1}$$

which depend on the nuclear and electronic coordinates, X and x, and nuclear coordinates X, respectively. Notice that the electronic Hamiltonian operator contains the nuclear repulsion terms.

It is also known, in the Born-Oppenheimer approximation, that the eigenfunctions of the full Hamiltonian operator (1) may be factorized into an electronic wave-function and a nuclear one:

$$\Phi^{n,\nu}(X,x) = \Psi_e^n(X,x)\phi_N^{n,\nu}(X) \tag{2}$$

In addition, since the electrons move much faster than the nuclei, and are much more energetic, it is possible to solve independently the electronic Schrödinger equation, assuming for the nuclei a frozen configuration, X_o:

$$H_e(X_o,x)\Psi_e^n(X_o,x) = E_{X_o}^n\Psi_e^n(X_o,x) \tag{3}$$

A set of electronic energy values, $E_{X_o}^n$, can then be obtained for various electronic states $\psi_e^n(X_o,x)$ characterized by the quantum number n. As a

matter of fact, the electronic energy does not depend explicitly on the nuclear coordinates, X_o. It is indeed a *numerical* value whose magnitude depends parametrically on the nuclear configuration considered. The dependence of $E^n_{X_o}$, versus the nuclear coordinates X, however, can be expressed explicitely by fitting conveniently on a set of $E^n_{X_i}$ values obtained solving equation (3) for different X_i.

$$\{E^n_{X_i}\} \rightarrow V^n(X) \tag{4}$$

Let us now turn on the full Schrödinger equation:

$$H\Phi^{n,v}(X,x) = [H_e(X,x) + T(X)]\Psi^n_e(X,x)\varphi^{n,v}_N(X) \tag{5}$$

and introduce equation (3) in it. After substituting $E^n_{X_i}$ by $V^n(X)$, we obtain:

$$[T(X) + H_e(X,x)]\psi^n_e(X,x)\varphi^{n,v}_N(X) \approx [T(X) + V^n(X)]\psi^n_e(X,x)\varphi^n_N(X) \tag{6}$$

The right member of this equation contains an operator which depends only on the nuclear coordinates, X, and on the quantum number n. As a result, the following approximation may be written for an effective nuclear Hamiltonian operator:

$$[T(X) + V^n(X)] = H^n_N(X) \tag{7}$$

which will depend on the electronic state n. Notice that $V^n(X)$ is an effective potential, which can exhibit eventually the same value for different X's.

The Group Theory for Non-Rigid Molecules considers isoenergetic isomers, and the interconversion motions between them. Because of the discernability between the identical nuclei, each isomer possesses a different electronic Hamiltonian operator in (3), with different eigenfunctions, but the same eigenvalue. In contrast, a non-rigid molecule has an unique effective nuclear Hamiltonian operator (7).

Usually there are many possible interconversion movements between these isomers, which can change one degenerate configuration into another one of same energy. These interconversion movements are internal rotations, inversions, ring puckerings, etc...They can be described by some operators, M_i, acting on X, which will leave the nuclear Hamiltonian operator (7) invariant.

Let us now consider the complete set of operators M_i. It will contain the identity operator, E, which maintains the unchanged configuration. On the other hand, if the set contains the operator M_i, which transforms one configuration into another, its inverse, M_i^{-1}, which transforms the former into the initial form, exists necessarily since we are considering non-rigid molecules in thermodynamic equilibrium.

It is easily seen that the complete set of operators M_i, which leave the nuclear Hamiltonian invariant (7), forms a group [21]. Indeed, if $[H, M_i] = 0$ and $[H, M_j] = 0$, the product M_iM_j will also commute:

$$[H, M_iM_j] = (H, M_iM_j - M_iM_j, H) =$$
$$= (H, M_iM_j - M_iHM_j) = [H, M_i]M_j \equiv 0 \tag{8}$$

In this way, it is shown that the product M_iM_j belongs also to the set of M_i.

Likewise it is easy to see that, if M_i commutes, its inverse M_i^{-1} will also commute:

$$0 \equiv [E, H] = [M_iM_i^{-1}, H] = (M_iM_i^{-1}H - HM_iM_i^{-1}) =$$
$$= M_i(M_i^{-1}H - HM_i^{-1}) = M_i[M_i^{-1}, H] \tag{9}$$

It may be then concluded that the set of the operations M_i, which commute with H, is complete, i.e., it forms a group, whenever the multiplication operations in (8) and (9) are associative. The order of the group will of course depend on the definition of Hamiltonian operator, which may be approximate. Resorting to this definition, we shall analyze the Longuet-Higgins and Altmann's theories, and we will propose a more general one, we call the **Non-Rigid Molecule Group Theory** (NRG). It has to be stated that some attempts to unify both points-of-view have be already done by Gilles et al. [25] and Erza [26], and that other theories exist, such as the **Isometric Group Theory** of Günthard [27]. Localized Hamiltonian operators have been considered by Natanson into the formalism of the Permutation-Inversion Groups [28].

2.1.1. Direct and semi-direct products

Throughout the present paper, the concepts of direct and semi-direct products will be used. Let us define them next. For this purpose, let us consider two groups, I and G, which have no element in common except for the identity.

If I and G are such that any element I_s commutes with any element G_r:

$$I_sG_r = G_rI_s \tag{10}$$

their product forms a group S which is the direct product of I and G:

$$I \times G = S \tag{11}$$

On the other hand, if I and G do not commute in detail but rather:

$$IG_r = G_rI$$

for all $G_r \in G$, then the group forms a set which is called the semi-direct product of I and G, and is expressed as:

$$I \wedge G = S \tag{12}$$

where the invariant subgroup I with respect to G is written on left side [29].

2.2. Longuet-Higgins' Theory

Longuet-Higgins observed that, if we consider a molecule as an ensemble of nuclei and electrons, the Hamiltonian operator of this molecule must be invariant (in absence of external perturbation) with respect to the following operations [5]:

a) Any permutation of the positions and spins into any set of identical particles;

b) Any rotation of the positions and spins of the whole ensemble of the particles around the center of mass;

c) Any translation of the same ensemble;

d) The reversal of all the angular and spin momenta;

e) The simultaneous inversion of the positions of all the particles with respect to the center of mass.

The complete group of the Hamiltonian operator is then the direct product of various groups. To study most of the non-rigid molecules, however, not all of the complete group needs to be taken into account. Some operations can occur during a short time (electron permutations) or during a too long time (optic isomer racemization) to be observed at the laboratory time scale. Therefore, Longuet-Higgins considers only the set of operations restricted to the *feasible* movements, i.e., those movements which can be performed without passing over an insuperable energy barrier. The **Molecular Symmetry Group** is then composed of all feasible operations.

Let P be any permutation of positions and spins of identical nuclei, or any product of such permutations. Let E be the identity, E^* the inversion of all particle positions with respect to the center of mass, and P^* the product $PE^* = E^*P$. Then the Molecular Symmetry Group is the set of:

a) All feasible P, including E,

b) All feasible P^*, not necessarily including E^*.

As an example of feasible operations, we have the ammonia inversion, phenomenon which has been observed experimentally.

Fig. 1-a Ammonia inversion as feasible operation.

As an example of non feasible operations, we have the methylfluoride in-

version, phenomenon which has not been observed experimentally.

Fig. 1-b Methylfluoride inversion as non-feasible operation.

Although the methyl group inversion appears to be analogue to that of ammonia, there exits a fluorine atom, which has to pass through the carbon atom, which is indeed energertically prohibited.

In order to illustrate the Molecular symmetry Group Theory, let us consider the methyl boron difluoride molecule ($CH_3 - BF_2$), which contains a nearly free rotating methyl group. We shall see next that the torsional levels of this molecule can be classified according to the irreducible representations of the symmetry point group C_{6v}, although this molecule does not possess, in a random configuration, any symmetry at all.

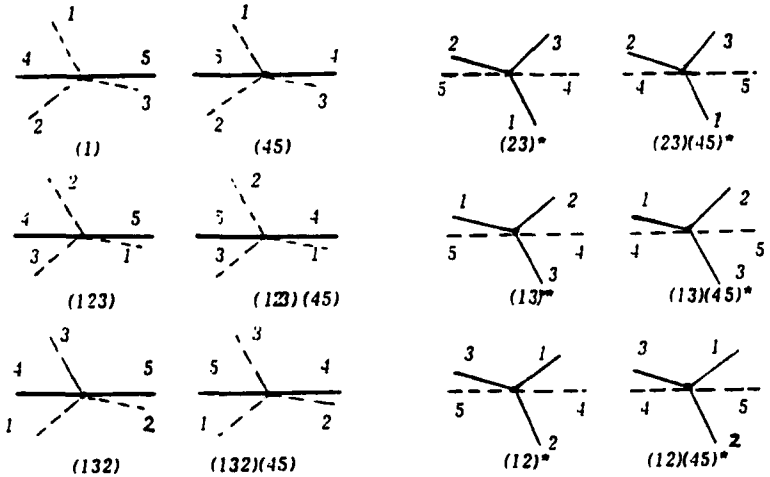


Fig.2. The twelve isoenergetic configurations and non-rigid operations corresponding to the internal rotation in methyl boron difluoride.

In Fig. 2, the twelve possible isoenergetic configurations that can be obtained by permutations and permutations-inversions of the hydrogen and fluorine atoms, are presented. In this figure, the identical particles, the protons

and the fluorine nuclei, are labelled from 1 to 3, and from 4 to 5, respectively, counterclock-wise. The six feasible permutations, E, (123), (132), (45), (123)(45) and (132)(45) are in fact internal rotations of the $-CH_3$ radical with respect to the reference plane $-BF_2$, of $0°$, $\frac{2\pi}{3}$, $\frac{4\pi}{3}$, π, $\frac{\pi}{3}$, and $\frac{5\pi}{3}$, respectively. The six feasible permutations-inversions (12)*, (13)*, (23)*, (12)(45)*, (13)(45)*, and (23)(45)*, which remove the fluorine atoms to the back of the figure, are equivalent to a rotation of the methyl radical to a symmetric position with respect to the plane $-BF_2$, after the above rotations. The existence of this kind of operations is related to the existence of symmetry planes in both the rotor and the reference frame [30].

These twelve operations form a group, called G_{12}. Its multiplication table, as well as, its character table are easily deduced. We present this in Table 1. The **Molecular Symmetry Group** of $CH_3 - BF_2$ is then seen to be isomorphic to the symmetry point group C_{6v}.

$$G_{12} \sim C_{6v} \tag{13}$$

Table 1: Character Table for the Molecular Symmetry Group G_{12} corresponding to the $BF_2 - CH_3$ molecule.

Rep.	E	(123) (132)	(12)* (13)* (23)*	(45)	(123)(45) (132)(45)	(12)(45)* (13)(45)* (23)(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

2.3. The Altmann's Theory

Beside the Longuet Higgins'theory, Altmann proposed another theory, generally equivalent and more understandable from the chemical-physical point of view. This theory is based on the existence of a general group, called **Schrödinger Supergroup**, equivalent to the **Molecular Symmetry Group** of Longuet-Higgins.

This general group could be factorized into two subgroups [10]:

$$S = I \wedge G \tag{12}$$

where the caret means semi-direct product, being I the invariant subgroup of the isodynamic operations, and G the Schrödinger subgroup of the symmetry operations of the molecule in a fixed random configuration [11,29].

In order to define these two subgroups Altmann wrote down the electronic Hamiltonian operator of the molecule under study in terms of the coordinates, r_i, and masses of the nuclei, m_i,

$$H_{X,Y,Z,x,y,z}(r_1, r_2, ...m_1, m_2) \tag{14}$$

The coordinates are expressed in the molecular axes x,y,z, which are rigidly attached to the molecule. These coordinates and masses are labelled in some laboratory axes, X,Y,Z, fixed in space. Under this definition, a symmetry operation is a change of axes that leaves the Hamiltonian operator (14) invariant, and the group of all such operations is the **Schrödinger subgroup.**

Altmann considered two types of operations that belong to the Schrödinger subgroup: the Euclidean and the discrete symmetry operations. Euclidean operations are those that change the laboratory axes, leaving the Hamiltonian operator invariant. They are translations and rotations of the whole molecule, in free space, in which the x,y,z molecular axes are kept constant. A discrete symmetry operation is a change of the molecular axes in such away as to induce permutations of the coordinates of identical particles [10].

Altmann remarked that, in free space, the Euclidean operations are not of physical interest. Therefore the Euclidean operations will be assimilated with the identity. Besides, he stated that the discrete symmetry operations are purely changes of labelling, especifically they are not motions of atoms. The Schrödinger subgroup may be then assimilated with the symmetry point group of the molecule in a fixed configuration.

The **isodynamic operation subgroup,** I, is then defined as the subgroup formed by all the interconversion operations, I_i, which transform a configuration into another isoenergetic one, and having the same symmetry properties, i.e., the isodynamic operations leave the nuclear Hamiltonian (7) invariant. These operations are internal rotations, inversions, ring puckerings, etc...The isodynamic subgroup is easily seen to form an invariant subset with respect to the symmetry point group operations [10].

Altmann emphasized that whereas the symmetry operations are simple changes of labels, the isodynamic operations describe material motions of group of atoms with respect to the rest of the molecule, which carry the labels with them, and the requirement of feasibility can be imposed on them.

As an example of application of the Altmann's theory, let us consider again the $CH_3 - BF_2$ molecule of Fig. 2. In this figure, the permutations of order three may be regarded as threefold rotations of the methyl moiety with respect

to the plane BF_2 around the C-B bond, and the permutations of order two as twofold rotations of the BF_2 moiety with respect to the same plane. In the same way, the permutations- inversions may be regarded as a rotation of the methyl moiety from θ to $-\theta$ with respect to the BF_2 plane, followed by the above operations (including E). This last operation is defined by the operator:

$$Uf(\theta) \equiv f(-\theta) \qquad (15-a)$$

which is called **switch**. The existence of this operation, as well as that of the permutation-inversion $(23)^*$, is conditioned by the presence of symmetry planes in both the rotor and the reference frame [30]. The switch subgroup is written as:

$$U_\theta^I = [E + U] \qquad (15-b)$$

The rotation subgroups can be written as:

$$C_3^I = [E + C_3 + C_3^2] \qquad (16-a)$$

and

$$C_2^I = [E + C_2] \qquad (16-b)$$

where C_3, C_3^2, and C_2 are rotation operations of $\frac{2\pi}{3}$, $\frac{4\pi}{3}$, and π, respectively. As a result, the isodynamic subgroup, I, may be expressed as the semi-direct product:

$$I = [C_3^I \times C_2^I] \wedge U_\theta^I \qquad (17-a)$$

Taking into account that the $CH_3 - BF_2$ molecule, in a random configuration, does not possess any symmetry at all, its Schrödinger subgroup, G, is the identity. The Schrödinger supergroup, S, of this molecule takes the form:

$$S = I \wedge E = [C_3^I \times C_2^I] \wedge U_\theta^I \sim C_{6v} \qquad (17-b)$$

which is a group of order twelve, isomorphic with the C_{6v} point group.

Inspection of Fig. 2, shows that the equivalence between the Schrödinger Supergroup and the Molecular Symmetry Group operations, for the CH_3-BF_2 molecule, is easily verified. In Table 2, the two groups appear to be identical.

2.4. The Non–Rigid Molecule Group Theory (NRG)

Group Theory for non-rigid molecules considers only large amplitude mouvements ignoring the small amplitude motions, such as vibrations. In the following, the **Non-Rigid Molecule Group** (NRG) will be strictly defined as the complete set of the *molecular* conversion operations, which commute

Table 2: Equivalence between the Molecular Symmetry Group and Schrödinger Supergroup operations in the $CH_3 - BF_2$ molecule.

$(1) = E$	$(45) = C_2$
$(123) = C_3$	$(45)(132) = C_6$
$(132) = C_3^2$	$(45)(123) = C_6^5$
$(23)^* = U$	$(45)(23)^* = U\ C_2$
$(12)^* = U\ C_3$	$(45)(13)^* = U\ C_6$
$(13)^* = U\ C_3^2$	$(45)(12)^* = U\ C_6^5$

with a given nuclear Hamiltonian operator such as (7), limited to large amplitude motions. In addition, these molecular conversion operations will be expressed in terms of physical operations, such as rotations, internal rotations, inversions, similarly as in the Altmann's theory, rather than in terms of permutations and permutations-inversions. This way of expressing the non-rigid operations is indeed more descriptive and flexible [21].

Such a group is strictly equivalent to the Longuet-Higgins' Molecular Symmetry Group, provided that the motions considered in the Hamiltonian operator are those described by the Longuet-Higgins' Group.

2.4.1. The full Non-Rigid Molecule Group

To clear up this definition let us consider a molecule in absence of any external perturbation, so its nuclear Hamiltonian operator can be written in terms of relative coordinates with respect to the center of mass, neglecting the translation coordinates. In this conditions, the nuclear Hamiltonian operator is written as:

$$H = T(X_e) + T(X_i) + T(X_e, X_i) + V(X_i) \tag{18}$$

In this expression $T(X_e)$ and $T(X_i)$ are the kinetic operators corresponding to the overall rotation and the intramolecular motions respectively. $T(X_e, X_i)$ is the coupling term between these two kinds of motions, and $V(X_i)$ the potential energy operator.

The complete set of the *molecular* conversion operations which commute with such an Hamiltonian operator (18) will contain overall rotation operations describing the molecule rotating as a whole, and internal motion operations describing molecular moieties moving with respect to the rest of the molecule. Such a group is called the **full Non-Rigid Molecule Group** (full NRG).

2.4.2. The restricted Non-Rigid Molecule Group

Let us remark that generally the coupling term between the external rotation and the internal motions is not very significant, so that it may be neglected in equation (18), at least in first approximation:

$$H^L = T(X_e) + [T(X_i) + V(X_i)] \tag{19}$$

In such a case, the overall rotation variables, which may be expressed as rotational angles around some orthogonal axes, and the internal motion variables, which may be written as internal coordinates, are completely separable. Because of this separability, this approximate full Hamiltonian operator may be regarded as *local*.

The complete set of the molecular conversion operations which commute with this approximate Hamiltonian operator (19) will define another group, we call the *local* full NRG. This new group may be larger than the *exact* full NRG group.

On the other hand, the complete set of the overall rotation operations which transform one rotational state into another isoenergetic one, i.e., which commute with the external rotation Hamiltonian operator, $T(X_e)$, will form a group called the **external rotation symmetry group**. Besides, the complete set of the intramolecular operations, which commute with the Hamiltonian operator restricted to the internal motions, $T(X_i) + V(X_i)$, will define another group, we call the **restricted Non-Rigid Molecule Group** (rNRG). Both groups commute since the external and internal coordinates are now completely separable.

Notice that, since the external rotation is relatively slower than the internal motions, the external rotation symmetry group may be expected to be isomorphic with the symmetry point group of the molecule in its most symmetric configuration [4]. As a result, the *local* full NRG, defined by operator (19), may be expected to be isomorphic to the direct product of the restricted NRG by the symmetry point group of the molecule:

$$G^L_{full\ NRG} \sim G_{restricted\ NRG} \times G_{point\ group} \tag{20}$$

similarly as in the Altmann's theory of the Schrödinger Supergroup (12). Let us remark that this similarity is only apparent. Besides, because of the particular Altmann's definition of the symmetry operations which do not describe any motion but are simple changes of label, the Schrödinger subgroup do not commute necessarily with the Isodynamic one [10].

In order to go farther into the Non-Rigid Molecule Theory, let us analyze expression (20) and compare the full and restrictred non-rigid groups. For this purpose, let us consider two border-line cases:

a) The non-rigid molecule in a random configuration has no symmetry;
b) The non-rigid molecule in a random configuration retains some symmetry.

a) When the molecule does not possess any symmetry in a random configuration, the full and restricted NRG's are seen to be isomorphic. To verify this affirmation let us turn for a moment to the permutation and permutation-inversion formalism and express both groups in that language. Since the external rotations in a asymmetric system cannot be described in terms of permutations and permutations-inversions of identical particules (its external group is the identity), both groups will have necessarily the same form in the longuet-Higgins formalism. As a result, both groups will have at least the same group structure in the physical operation formalism, i.e., they are isomorphic.

When the non-rigid system has no symmetry in any configuration, the symmetry point group in (20) will be the identity, and the full NRG will coincide with the restricted one. Furthermore, when the non-rigid system does not have symmetry in a random configuration, but retains only some symmetry in particular configurations, the symmetry point group in (20) will be that of the molecule in its most symmetric configuration, and the restricted NRG will be isomorphic to the full NRG. In this case, the *local* full NRG will be larger than the *exact* full one.

b) When the molecule retains some symmetry in a random configuration, the symmetry point group in (20) will be also that of the system in its most symmetric configuration, and the restricted NRG group will be smaller than the full NRG.

One of the difficulties of the NRG theory is to construct correctly the full, as well as the restricted Hamiltonian operators, and to deduce properly the physical operations which commute with these operators. In many cases, however, the interconversion operations may be described easily as rotations of molecular moieties around some axes supported by a solid molecular frame. In such cases, the concept of restricted NRG recovers special relevance, and the restricted NRG appears to be equivalent to the isometric group [27].

When the separation of the external and internal motions is not possible, equation (20) does not necessarily hold, and the concept of restricted non-rigid molecule group vanishes. This is especially true in the case of molecules in which the reference frame is a single atom. In this case indeed the internal motions cannot be separated from the external motions (see section 6.2.).

In the two next sections, we shall consider molecules of increasing complexity, the non-rigidity of which proceeds from internal rotations, and we shall deduce their restricted NRG. Later on, restricted and full NRG's will be considered.

3. NON RIGID MOLECULE GROUPS FOR SYSTEMS WITH ONE, TWO AND THREE DEGREES OF FREEDOM

Up to now we have considered simple molecules with only one degree of internal freedom, such as $CH_3 - BF_2$, and we have deduced their Molecular Symmetry Groups or their Schrödinger Supergroups. In the following, we will introduce the Non Rigid Molecule Group Theory. For this purpose we shall consider non-rigid molecules of increasing complexity with one, two or three degrees of freedom [21]. In such a way, we will introduce in a gradual manner, the mathematical machinery useful for applying the NRG theory to problems of physical-chemical interest.

3.1. Phenol

The phenol molecule has only one rotor: its hydroxyl moiety which is able to rotate around the C_2 symmetry axis of the phenyl ring.

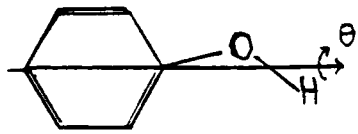


Fig. 3. Internal rotation in phenol.

The restricted Hamiltonian operator describing such a motion may be written as [31,32]:

$$H = [-\frac{\partial}{\partial\theta}B(\theta)\frac{\partial}{\partial\theta} + V(\theta)] \tag{21}$$

where $B(\theta)$ is the internal rotation constant, which depends mildly on the rotation angle θ and $V(\theta)$ is the potential energy operator in which the rotor is moving.

Normally, $B(\theta)$ and $V(\theta)$ are structural parameters having the same symmetry properties. Therefore, we regard $B(\theta)$ as a constant and limit our considerations to the potential energy function. In order to deduce the form of $V(\theta)$, let us note:

a) The twofold periodicity of the rotation;

b) The invariance of the energy with respect to the sense of the rotation.

As a result, a two fold energy function, which excludes any antisymmetric function (sine function) may be written for (21):

$$H = [-B\frac{\partial^2}{\partial\theta^2} + \sum_K A_K^C \cos 2K\theta] \tag{22}$$

The restricted NRG for the internal rotation in phenol may be then written as the sets of the internal motion operations which leave (22) invariant. The following group is easily deduced:

$$G_{rNRG} = [C_2^I \times U^I] \sim C_{2v} \tag{23}$$

which is a group of order four, isomorphic to the symmetry point group C_{2v}. In this expression, U^I is the subgroup of the simple switch (15-b), and C_2^I is the subgroup of the twofold rotation (16-b). The character table of such a group is written in Table 3.

Table 3: Character table for the internal rotation in phenol.

	E	C_2	U	UC_2
A_1	1	1	1	1
A_2	1	1	-1	-1
B_1	1	-1	1	-1
B_2	1	-1	-1	1

From this table, the symmetry eigenvectors which factorize the matrix Hamiltonian corresponding to operator (22) into boxes are easily deduced in the base of the solutions of the planar free rotor equation. There are:

$$\begin{array}{ll} \chi_{A_1} = \cos 2K\theta & \chi_{B_1} = \cos(2K+1)\theta \\ \chi_{A_2} = \sin 2K\theta & \chi_{B_2} = \sin(2K+1)\theta \end{array} \tag{24}$$

Since phenol in a random conformation has no symmetry, its full NRG is expected to be isomorphic to the restricted one.

3.2. Benzaldehyde

Let us now consider a molecule which has a second vibrational degree of freedom in addition to a rotor of order two. Such a molecule could be benzaldehyde in which an out-of-plane wagging mode of the aldehydic hydrogen atom is allowed. See figure 4.

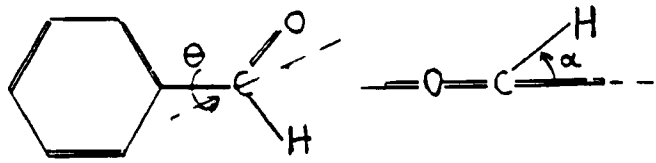


Figure 4. The benzaldehyde molecule in standard and stereographic projections.

In this figure, θ is the rotation angle of the phenyl moiety. The out-of-plane vibration is described by a torsion angle, α, around a perpendicular axis to the C-H bond and lying in the plane of the molecule.

The restricted Hamiltonian operator corresponding to these motions may be easily deduced if we consider:

a) The two-fold rotation of the phenyl group;

b) The one-fold rotation of the aldehydic hydrogen atom:

c) The invariance of the energy under the simultaneous change of the sense of the rotations. This last properties is closely related with the existence of symmetry planes in the rotors and the frame [33].

As a result, the restricted Hamiltonian operator is written as:

$$H = [-B_{11}\frac{\partial^2}{\partial\theta^2} - 2B_{12}\frac{\partial^2}{\partial\theta\partial\alpha} - B_{22}\frac{\partial^2}{\partial\alpha^2}] + \\ + \sum_K \sum_L [A_{KL}^{CC} \cos 2K\theta \cos L\alpha + A_{KL}^{SS} \sin 2K\theta \sin L\alpha] \qquad (25)$$

where the potential energy terms does not contain any antisymmetric products of trigonometric functions.

The restricted NRG for the internal rotation and wagging in benzaldehyde may be then written as the set of internal motion operators which leave the Hamiltonian operator (25) invariant:

$$G_{rNRG} = [C_2^I \times V^I] \sim C_{2v} \qquad (26)$$

which is another group of order two, isomorphic to the previous (23) and to the symmetry point group C_{2v}. In this new group, a new subgroup, V^I, called double switch, appears:

$$V^I = [E + V] \qquad (27)$$

This contains the identity and the double switch operations. The double switch operation, introduced into the formalism of Altmann [10], is defined as a double simultaneous rotation from θ to $-\theta$, and from α to $-\alpha$:

$$Vf(\theta,\alpha) \equiv f(-\theta,-\alpha) \tag{28}$$

As in the case of the simple switch, the existence of this operation is conditioned by the presence of symmetry planes in the rotors and the frame [33]. The character table of such a group may be written as shown in Table 4.

Table 4: Character table for the internal rotation and wagging in benzaldehyde.

	E	C_2	V	VC_2
A_1	1	1	1	1
A_2	1	1	-1	-1
B_1	1	-1	1	-1
B_2	1	-1	-1	1

From this table the symmetry eigenvectors are easily deduced in the base of the solutions of the double free rotor equation. They factorize into boxes the Hamiltonian matrix corresponding to operator (26).

These have the form:

$$\chi_{A_1} = \begin{cases} \cos 2K\theta \cos L\alpha \\ \sin 2K\theta \sin L\alpha \end{cases} \quad \chi_{A_2} = \begin{cases} \cos 2K\theta \sin L\alpha \\ \sin 2K\theta \cos L\alpha \end{cases}$$

$$\chi_{B_1} = \begin{cases} \cos(2K+1)\theta \cos L\alpha \\ \sin(2K+1)\theta \sin L\alpha \end{cases} \quad \chi_{B_2} = \begin{cases} \cos(2K+1)\theta \sin L\alpha \\ \sin(2K+1)\theta \cos L\alpha \end{cases} \tag{29}$$

where the rotational functions of periodicity even and odd correspond respectively to the A_n and B_n representations as in phenol (24).

3.3. Pyrocatechin

Let us now consider a molecule, such as pyrocatechin, which possesses two onefold rotors equivalent in structure and location. The rotational coordinates of such a molecular system are θ_1 and θ_2 (see Figure 5).

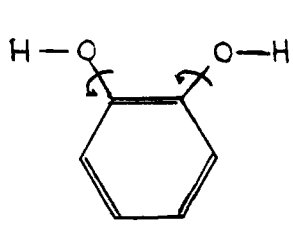


Figure 5: The double equivalent internal rotation in pyrocatechin.

The restricted Hamiltonian operator corresponding to these internal rotations may be deduced if we consider:

a) The onefold rotation of each rotor;

b) The equivalence of the rotors;

c) The invariance of the energy with respect to a simultaneous change of the sense of the rotations.

As a result, the Hamiltonian operator may be written as:

$$\begin{aligned} H = & [-B_1\frac{\partial^2}{\partial\theta_1^2} - 2B_{12}\frac{\partial^2}{\partial\theta_1\partial\theta_2} - B_2\frac{\partial^2}{\partial\theta_2^2}]+ \\ & +\sum_K\sum_L[A_{KL}^{CC}(cosK\theta_1 cosL\theta_2 + cosL\theta_1 cosK\theta_2)+ \\ & +[A_{KL}^{SS}(sinK\theta_1 sinL\theta_2 + sinL\theta_1 sinK\theta_2)] \end{aligned} \tag{30}$$

where the potential energy terms are symmetric with respect to the exchange and a simultaneous sign change of the rotation angles.

The restricted NRG for the double equivalent internal rotation in pyrocatechin may be then expressed as:

$$G_{rNRG} = [C_i^I \times C_{i'}^I] \times [W^I \times V^I] \sim C_{2v} \tag{31}$$

which is again a group of order two isomorphic to (23), (26) and the symmetry point group C_{2v}. In this group, V^I is the double switch subgroup, W^I is a new subgroup called the exchange subgroup, defined as:

$$W^I = [E + W] \tag{32}$$

This subgroup is constructed by the identity and angle exchange operation, which was also introduced into the formalism of Altmann. This was defined as [10]:

$$Wf(\theta_1, \theta_2) \equiv f(\theta_2, \theta_1) \tag{33}$$

The existence of this operation is conditioned by the equivalence of the rotors by superposition [33]. The character table of this group is given in Table 5.

Table 5: Character table for the double equivalent internal rotation of pyrocatechin.

	E	WV	W	V
A_1	1	1	1	1
A_2	1	1	-1	-1
A_3	1	-1	1	-1
A_4	1	-1	-1	1

Notice that in this table we retained the Longuet-Higgins' notation for the irreducible representations A_n in accordance to that of ethane [4].

From this character table, the symmetry eigenvectors for the double internal rotation in planar pyrocatechin are easily deduced on the basis of the double free rotor equation solutions:

$$\chi_{A_1} = \begin{cases} \cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2 \\ \sin K\theta_1 \sin L\theta_2 + \sin L\theta_1 \sin K\theta_2 \end{cases}$$

$$\chi_{A_2} = \cos K\theta_1 \sin L\theta_2 - \sin L\theta_1 \cos K\theta_2$$

$$\chi_{A_3} = \cos K\theta_1 \sin L\theta_2 + \sin L\theta_1 \cos K\theta_2 \qquad (34)$$

$$\chi_{A_4} = \begin{cases} \cos K\theta_1 \cos L\theta_2 - \cos L\theta_1 \cos K\theta_2 \\ \sin K\theta_1 \sin L\theta_2 - \sin L\theta_1 \sin K\theta_2 \end{cases}$$

The equivalence of the rotors is seen to induce binomial forms, symmetric or antisymmetric with respect to the angle exchange, whereas the symmetry planes are seen to induce symmetric or antisymmetric functions with respect to the plane of the molecule, i.e., a simultaneous sign change of the rotation angles, being A_2 the completely antisymmetric representation.

The *sine* × *sine* symmetric binomial forms are considered to introduce a cog-wheel effect between the rotors into the Hamiltonian operator (30), because each *sine* function depends on the sense of the rotation [34,35].

As in phenol and benzaldehyde, pyrocatechin in a random configuration has no symmetry, and thus the full NRG is expected to be isomorphic with the restricted one.

3.4. Pyrocatechin with Wagging Vibration Mode

Let us consider again the pyrocatechin molecule, but in a non–planar configuration, where the hydrogen atoms are in meta positions wagging out of plane synchronously (see Fig. 6).

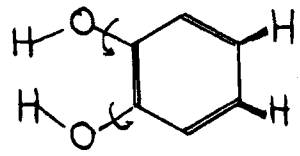


Figure 6. Double equivalent rotation and wagging mode in non–planar pyrocatechin.

The restricted Hamiltonian operator describing such nuclear motions (torsions + wagging) is a function of three variables, the torsion angles, θ_1 and θ_2, and the wagging angle, α. It is not so easily transcribed as in the previous cases. It may be expressed as a sum of kinetic operators, (the second partial derivatives with respect to all the variables) plus the potential energy function, which may be written as an expansion in terms of the complety symmetric eigenvectors. Therefore, the rNRG should be more conveniently deduced before.

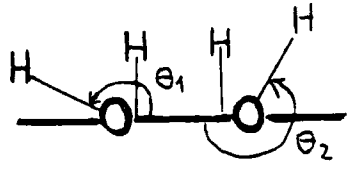
Due to the existence of the out–of–plane wagging, the two rotors are not equivalent anymore by simple superposition. They appear to be, however. equivalent by reflection in a plane perpendicular to the molecular plane dividing the molecule in two equal moeities.

In fact, it is easily seen that there are four isoenergetic configurations, as illustrated in Figure 7.

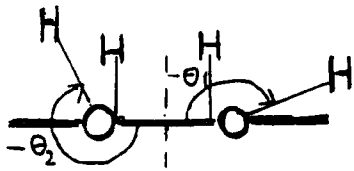
In this figure, the non–planar pyrocatechin molecule has been represented in stereographic projection, with the hydroxyl groups rotated in a random way by angles θ_1 and θ_2, and the two hydrogen atoms deviated at angle α from the plane of the molecule, configuration (1).

The restricted NRG can be easily deduced by taking into account that the enantiomorphic configuration (2), obtained by reflection in the molecular plane is isoenergetic. The corresponding operator is the triple switch operator:

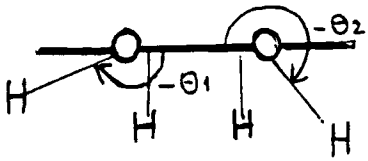
$$(VU)f(\theta_1, \theta_2, \alpha) \equiv f(-\theta_1, -\theta_2, -\alpha) \tag{35}$$



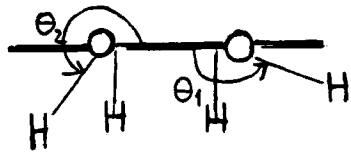
Configuration (1).



Configuration (3).



Configuration (2).



Configuration (4).

Figure 7. The four isoenergetic configurations of non-planar pyrocatechin in stereographic projection.

In the same way, it is clear that the configuration (3), obtained by reflection in a plane perpendicular to the molecular plane is isoenergetic., i.e., the two rotors are equivalent by reflection. The operator corresponding to this internal motion is the double–switch–exchange operator:

$$(WV)f(\theta_1, \theta_2, \alpha) \equiv f(-\theta_2, -\theta_1, \alpha) \tag{36}$$

Finally, the configuration (4) obtained by reflection in the molecular plane from the former (3) is also isoenergetic. The operator of such an internal motion is the simple–switch–exchange operator:

$$(WU)f(\theta_1, \theta_2, \alpha) \equiv f(\theta_2, \theta_1, -\alpha) \tag{37}$$

The non–rigid subgroup corresponding to these interconversion operators is:

$$G = [E + UV + WV + WU] = [(WV)^I \times (VU)^I] \tag{38}$$

and the restricted non–rigid group:

$$G_{rNRG} = [C_1^I \times C_{1'}^I] \wedge [(WV)^I \times (VU)^I] \sim C_{2v} \tag{39}$$

which is again a group of order 4 isomorphic to the symmetry point group C_{2v}. The character table of such a group is given in Table 6.

From this character table, the symmetry eigenvectors are easily deduced in the basis of the solutions of the triple free rotor equation:

Table 6: Character table for the double rotation and wagging in non-planar pyrocatechin.

	E	WV	WU	VU
A_1	1	1	1	1
A_2	1	1	-1	-1
A_3	1	-1	1	-1
A_4	1	-1	-1	1

$$\chi_{A_1} = \begin{cases} (\cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2) \cos M\alpha \\ (\sin K\theta_1 \sin L\theta_2 + \sin L\theta_1 \sin K\theta_2) \cos M\alpha \\ (\cos K\theta_1 \sin L\theta_2 - \sin L\theta_1 \cos K\theta_2) \sin M\alpha \end{cases}$$

$$\chi_{A_2} = \begin{cases} (\cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2) \sin M\alpha \\ (\sin K\theta_1 \sin L\theta_2 + \sin L\theta_1 \sin K\theta_2) \sin M\alpha \\ (\cos K\theta_1 \sin L\theta_2 - \sin L\theta_1 \cos K\theta_2) \cos M\alpha \end{cases}$$

$$\chi_{A_3} = \begin{cases} (\cos K\theta_1 \cos L\theta_2 - \cos L\theta_1 \cos K\theta_2) \sin M\alpha \\ (\sin K\theta_1 \sin L\theta_2 - \sin L\theta_1 \sin K\theta_2) \sin M\alpha \\ (\cos K\theta_1 \sin L\theta_2 + \sin L\theta_1 \cos K\theta_2) \cos M\alpha \end{cases}$$

$$\chi_{A_4} = \begin{cases} (\cos K\theta_1 \cos L\theta_2 - \cos L\theta_1 \cos K\theta_2) \cos M\alpha \\ (\sin K\theta_1 \sin L\theta_2 - \sin L\theta_1 \sin K\theta_2) \cos M\alpha \\ (\cos K\theta_1 \sin L\theta_2 + \sin L\theta_1 \cos K\theta_2) \sin M\alpha \end{cases} \tag{40}$$

The restricted Hamiltonian operator for such nuclear motions may be expressed by writing the potential as a linear combination of the completely symmetric eigenvectors [33,34]:

$$V(\theta_1, \theta_2, \alpha) = \sum_K \sum_L \sum_M [A^{CCC}_{KLM}(\cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2) \cos M\alpha$$

$$+ A^{SSC}_{KLM}(\sin K\theta_1 \sin L\theta_2 + \sin L\theta_1 \sin K\theta_2) \cos M\alpha]$$

$$+ \sum_K \sum_L \sum_M A^{CSS}_{KLM}(\cos K\theta_1 \sin L\theta_2 - \sin L\theta_1 \cos K\theta_2) \sin M\alpha \tag{41}$$

In this expression, the cross products between cosine and sine functions, which depend on the sense of rotations introduce the cog-wheel effects between

the rotors and the wagging hydrogen atoms. When $\alpha = 0$, this expression is reduced to the potential (30) for the planar pyrocatechin molecule.

As in the previous case, the full NRG is expected to be isomorphic with the restricted one, because the molecule in a random configuration has no symmetry.

4. NON-RIGID MOLECULE GROUPS FOR C_{3v} ROTOR MOLECULES

In section 3, we considered non–rigid systems of increasing complexity in order to introduce the most essential interconversion operations. The groups which allow us to classify the energy levels were isomorphic to the symmetry point group C_{2v}. Next, we shall consider more complex molecules bearing a C_{3v} rotors. This simple increment of symmetry increases unexpectedly the complexity of the non–rigid groups [21].

4.1. Planar Acetone

We have already studied linear two rotor molecules such as $CH_3 - BF_2$ which possesses C_{3v} and C_{2v} rotors, but only with one degree of freedom. Let us now consider acetone-like molecules which possess a C_{2v} frame, such as pyrocatechin, and two C_{3v} rotors. The restricted Hamiltonian operator of such molecules is the same as that of pyrocatechin (32) except for the potential energy function which has to reflect the threefold periodicity of the rotors.

The restricted NRG for acetone can be is easily deduced from that of pyrocatechin (31) by introducing the threefold periodicity of the rotors:

$$G_{rNRG} = [C_3^I \times C_{3'}^I] \wedge [W^I \times V^I] = G_{36} \tag{42}$$

where C_3^I and $C_{3'}^I$ are the subgroup of the threefold rotation of the two rotors (16-a), and $[W^I \times V^I]$ is the subgroup corresponding to the C_{2v} frame (31).

The acetone-like G_{36} group is a group of order 36, that has no equivalent in the symmetry point groups. It is formally identical to the NRG of ethane [22] (see expression 67), although the operations do not have the same physical meaning. Similarly, the Molecular Symmetry Groups of both molecules are identical. In contrast, the Altmann theory gives two groups of same structure which are formally different [19] (see expression 71).

Table 7: Character table for the double internal rotation in planar acetone with interaction between the rotors.

	E	$2C_3$ $2C_{3'}$	2 $C_3C_{3'}$	2 $C_3C_{3'}^2$	W $2W\times$ $C_3C_{3'}^2$	$6W\times$ C_3C_3	WV $2WV\times$ $C_3C_{3'}$	$6WV\times$ $C_3C_{3'}$	V $8V\times$ $C_3C_{3'}$
A_1	1	1	1	1	1	1	1	1	1
A_2	1	1	1	1	-1	-1	1	1	-1
A_3	1	1	1	1	1	1	-1	-1	-1
A_4	1	1	1	1	-1	-1	-1	-1	1
E_1	2	-1	2	-1	0	0	2	-1	0
E_2	2	-1	2	-1	0	0	-2	1	0
E_3	2	-1	-1	2	2	-1	0	0	0
E_4	2	-1	-1	2	-2	1	0	0	0
G	4	1	-2	-2	0	0	0	0	0

From the multiplication table, the character table for the G_{36} for planar acetone-like molecule is easily deduced. This character table can be found in literature [8,19,21,22,34,37]. It is presented in Table 7. It presents nine irreducible representations, 4 non-degenerate, 4 twofold degenerate and 2 fourfold degenerate. Into the symmetry point groups, nor so large groups, nor four fold degenerate representations, are observed except for the I_h group.

The symmetry eigenvectors corresponding to the G_{36} acetone-like group are not so easily deduced by tryings and errors, from the character table, as in the case of the one-fold rotor molecules. The projector technique has to be used for [21,34]:

$$P_j = \frac{l_j}{N} \sum_R^N \chi_j(R) \tag{43}$$

In this expression, N is the order of the group, l_j the dimension of the irreducible representation j on which the operator projects, χ_j is the character of the representation for the operation R, being R the corresponding operator. The sum runs over all the operations of the group.

Applying this projector on each of the terms of a rotational configuration interaction expansion, built up on the basis of the double free rotor solutions, the symmetry eigenvectors are obtained [21,22,34].

For the non–degenerate irreducible representations we have:

$$A_1 \quad \begin{cases} \chi_{A_1}^{CC} = \cos 3K\theta_1 \cos 3L\theta_2 + \cos 3L\theta_1 \cos 3K\theta_2 \\ \chi_{A_1}^{SS} = \sin 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \sin 3K\theta_2 \end{cases} \tag{44}$$

where $L \geq K$ in order to avoid repetition.

$$A_2 \qquad \chi^{CS}_{A_2} = \cos 3K\theta_1 \sin 3L\theta_2 - \sin 3L\theta_1 \cos 3K\theta_2$$

$$A_3 \qquad \chi^{CS}_{A_3} = \cos 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \cos 3K\theta_2$$

where L and K may take all integer values.

$$A_4 \begin{cases} \chi^{CC}_{A_4} = \cos 3K\theta_1 \cos 3L\theta_2 - \cos 3L\theta_1 \cos 3K\theta_2 \\ \chi^{SS}_{A_4} = \sin 3K\theta_1 \sin 3L\theta_2 - \sin 3L\theta_1 \sin 3K\theta_2 \end{cases}$$

where $L > K$.

For the twofold degenerate irreducible representations we have:

$$E_1 \begin{cases} \chi^{CC}_{E_1} = [\cos(3K \pm \delta)\theta_1 \cos(3L + \delta)\theta_2 \pm \sin(3K \pm \delta)\theta_1 \sin(3L + \delta)\theta_2 + \\ \quad + \cos(3L + \delta)\theta_1 \cos(3K \pm \delta)\theta_2 \pm \sin(3L + \delta)\theta_1 \sin(3K \pm \delta)\theta_2] \\ \chi^{CS}_{E_1} = [\cos(3K \pm \delta)\theta_1 \sin(3L \pm \delta)\theta_2 \mp \sin(3K \pm \delta)\theta_1 \cos(3L + \delta)\theta_2 - \\ \quad - \sin(3L + \delta)\theta_1 \cos(3K \pm \delta)\theta_2 \pm \cos(3L + \delta)\theta_1 \sin(3K \pm \delta)\theta_2] \end{cases}$$

$$E_2 \begin{cases} \chi^{CC}_{E_2} = [\cos(3K \pm \delta)\theta_1 \cos(3L + \delta)\theta_2 \pm \sin(3K \pm \delta)\theta_1 \sin(3L + \delta)\theta_2 - \\ \quad - \cos(3L + \delta)\theta_1 \cos(3K \pm \delta)\theta_2 \mp \sin(3L + \delta)\theta_1 \sin(3K \pm \delta)\theta_2] \\ \chi^{CS}_{E_2} = [\cos(3K \pm \delta)\theta_1 \sin(3L \pm \delta)\theta_2 \mp \sin(3K \pm \delta)\theta_1 \cos(3L + \delta)\theta_2 + \\ \quad + \sin(3L + \delta)\theta_1 \cos(3K \pm \delta)\theta_2 \mp \cos(3L + \delta)\theta_1 \sin(3K \pm \delta)\theta_2] \end{cases} \tag{45}$$

$$E_3 \begin{cases} \chi^{CC}_{E_3} = [\cos(3K \pm \delta)\theta_1 \cos(3L + \delta)\theta_2 \mp \sin(3K \pm \delta)\theta_1 \sin(3L + \delta)\theta_2 + \\ \quad + \cos(3L + \delta)\theta_1 \cos(3K \pm \delta)\theta_2 \mp \sin(3L + \delta)\theta_1 \sin(3K \pm \delta)\theta_2] \\ \chi^{CS}_{E_3} = [\cos(3K \pm \delta)\theta_1 \sin(3L \pm \delta)\theta_2 \pm \sin(3K \pm \delta)\theta_1 \cos(3L + \delta)\theta_2 + \\ \quad + \sin(3L + \delta)\theta_1 \cos(3K \pm \delta)\theta_2 \pm \cos(3L + \delta)\theta_1 \sin(3K \pm \delta)\theta_2] \end{cases}$$

$$E_4 \begin{cases} \chi^{CC}_{E_4} = [\cos(3K \pm \delta)\theta_1 \cos(3L + \delta)\theta_2 \mp \sin(3K \pm \delta)\theta_1 \sin(3L + \delta)\theta_2 - \\ \quad - \cos(3L + \delta)\theta_1 \cos(3K \pm \delta)\theta_2 \pm \sin(3L + \delta)\theta_1 \sin(3K \pm \delta)\theta_2] \\ \chi^{CS}_{E_4} = [\cos(3K \pm \delta)\theta_1 \sin(3L \pm \delta)\theta_2 \pm \sin(3K \pm \delta)\theta_1 \cos(3L + \delta)\theta_2 - \\ \quad - \sin(3L + \delta)\theta_1 \cos(3K \pm \delta)\theta_2 \mp \cos(3L + \delta)\theta_1 \sin(3K \pm \delta)\theta_2] \end{cases}$$

In all these expressions $\delta = \pm 1$ and $(3K \pm \delta)$ and $(3L + \delta)$ are always positive integers, $L \geq K$ for E_1 and E_3, and $L > K$ for E_2 and E_4.

For the fourfold degenerate irreducible representation we have:

$$G_1 \begin{cases} \chi^{CC}_{G_1} = \cos 3K\theta_1 \cos(3L + \delta)\theta_2 \\ \chi^{SS}_{G_1} = \sin 3K\theta_1 \sin(3L + \delta)\theta_2 \end{cases}$$

$$G_2 \begin{cases} \chi^{CS}_{G_2} = \cos 3K\theta_1 \sin(3L + \delta)\theta_2 \\ \chi^{SC}_{G_2} = \sin 3K\theta_1 \cos(3L + \delta)\theta_2 \end{cases} \tag{46}$$

$$G_3 \begin{cases} \chi^{CC}_{G_3} = \cos(3K + \delta)\theta_1 \cos 3L\theta_2 \\ \chi^{SS}_{G_3} = \sin(3K + \delta)\theta_2 \sin 3L\theta_2 \end{cases}$$

$$G_4 \begin{cases} \chi^{CS}_{G_4} = \cos(3K+\delta)\theta_1 \sin 3L\theta_2 \\ \chi^{SC}_{G_4} = \sin(3K+\delta)\theta_2 \cos 3L\theta_2 \end{cases}$$

where G_1, G_2,G_3 and G_4 are the four components of the irreducible representation, $\delta = \pm 1$, and $(3K+\delta)$ and $(3L+\delta)$ are always positive integers. K and L may have all the possible values.

The potential energy function of the Hamiltonian operator for acetone-like molecules may be written as a linear combination of all the A_1 eigenvectors [44]:

$$V(\theta_1,\theta_2) = \sum_K \sum_L [A^{CC}_{KL}(\cos 3K\theta_1 \cos 3L\theta_2 + \cos 3L\theta_1 \cos 3K\theta_2) + \\ + A^{SS}_{KL}(\sin 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \sin 3L\theta_2)] \qquad (47)$$

expression identical to that of pyrocatechin (30), except for the threefold periodicity.

4.2. Pyramidal Acetone (Dimethylamine)

Finally, let us consider once more the acetone molecule, but in a pyramidal configuration with the oxygen atom outside the CCC plane. Acetone exhibits this configuration in the lowest triplet state. In addition, pyramidal acetone may undergo an inversion phenomenon or wagging of the oxygen atom up and down the CCC plane.

The restricted Hamiltonian operator for the double C_{3v} rotation and wagging of the oxygen atom in pyramidal acetone is again a function of three variables the two torsion angles, θ_1 and θ_2, and the wagging angle α. This operator should be invariant under the triple switch operator, VU, the double switch–exchange operator, WV, and the single switch–exchange operator, and WU as, defined by (36), (37), and (38), as well as under the threefold rotation of each rotor. Dimethylamine is another example of molecules undergoing such motions.

As result, the restricted NRG for the pyramidal acetone is written as:

$$G_{rNRG} = (C_3^I \times C_{3'}^I) \wedge [(WV)^I \times (VU)^I] \qquad (48)$$

which is other group of order 36.

Notice that, in the subgroup

$$[(WV)^I \times (VU)^I] = [E + WV + WU + VU]$$

WU and VU possess exactly the same commutation properties with respect to the remaining group elements as W and V, because the operation U commutes with all of them. U acts indeed only on variable α. Therefore, the pyramidal acetone NRG (48) possesses the same group structure and the same character table as the NRG of acetone, just as it happens in planar and non–planar pyrocatechin. This is given in Table 8.

Table 8: Character Table for the double internal rotation and wagging in pyramidal acetone with interaction between the moving parts.

	E	$2C_3$ $2C_{3'}$	2 $C_3C_{3'}$	2 $C_3C_{3'}^2$	WU $2WU\times$ $C_3C_{3'}^2$	$6WU\times$ $C_3C_{3'}$	$2WV\times$ $C_3C_{3'}$	WV $6WV\times$ $C_3C_{3'}$	VU $8VU\times$ $C_3C_{3'}$
A_1	1	1	1	1	1	1	1	1	1
A_2	1	1	1	1	-1	-1	1	1	-1
A_3	1	1	1	1	1	1	-1	-1	-1
A_4	1	1	1	1	-1	-1	-1	-1	1
E_1	2	-1	2	-1	0	0	2	-1	0
E_2	2	-1	2	-1	0	0	-2	1	0
E_3	2	-1	-1	2	2	-1	0	0	0
E_4	2	-1	-1	2	-2	1	0	0	0
G	4	1	-2	-2	0	0	0	0	0

From this character table, and the symmetry eigenvectors of planar acetone (54-56), the symmetry eigenvectors of pyramidal acetone are easily deducible. For this purpose, linear combinations of the eigenvectors, which exhibit the same behavior for all the operations except for WU and VU, are built up. In addition, to the coefficients of which are trigonometric functions of the wagging angle, α. The coefficients are chosen in such a way that the linear combinations fulfill the characters corresponding to operators WU and VU.

For the non–degenerate irreducible representations, the following eigenvectors are encountered:

$$A_1 \begin{cases} \chi_{A_1}^{CCC} = [\cos 3K\theta_1 \cos 3L\theta_2 + \cos 3L\theta_1 \cos 3K\theta_2] \cos M\alpha \\ \chi_{A_1}^{SSC} = [\sin 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \sin 3K\theta_2] \cos M\alpha \\ \chi_{A_1}^{SCS} = [\cos 3K\theta_1 \sin 3L\theta_2 - \sin 3L\theta_1 \cos 3K\theta_2] \sin M\alpha \end{cases}$$

$$A_2 \begin{cases} \chi_{A_2}^{CCS} = [\cos 3K\theta_1 \cos 3L\theta_2 + \cos 3L\theta_1 \cos 3K\theta_2] \sin M\alpha \\ \chi_{A_2}^{SSS} = [\sin 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \sin 3K\theta_2] \sin M\alpha \\ \chi_{A_2}^{CSS} = [\cos 3K\theta_1 \sin 3L\theta_2 - \sin 3L\theta_1 \cos 3K\theta_2] \cos M\alpha \end{cases} \quad (49)$$

$$A_3 \begin{cases} \chi_{A_3}^{CCS} = [\cos 3K\theta_1 \cos 3L\theta_2 - \cos 3L\theta_1 \cos 3K\theta_2] \sin M\alpha \\ \chi_{A_3}^{SSS} = [\sin 3K\theta_1 \sin 3L\theta_2 - \sin 3L\theta_1 \sin 3K\theta_2] \sin M\alpha \\ \chi_{A_3}^{SCS} = [\cos 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \cos 3K\theta_2] \cos M\alpha \end{cases}$$

$$A_4 \begin{cases} \chi_{A_4}^{CCC} = [\cos 3K\theta_1 \cos 3L\theta_2 - \cos 3L\theta_1 \cos 3K\theta_2] \cos M\alpha \\ \chi_{A_4}^{SSC} = [\sin 3K\theta_1 \sin 3L\theta_2 - \sin 3L\theta_1 \sin 3K\theta_2] \cos M\alpha \\ \chi_{A_4}^{SCS} = [\cos 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \cos 3K\theta_2] \sin M\alpha \end{cases}$$

For the twofold degenerate irreducible representations, the following linear combinations are found:

$$E_1 \begin{cases} \chi_{E_1}^{CCC} = \chi_{E_1}^{CC} \cos M\alpha + \chi_{E_1}^{CS} \sin M\alpha \\ \chi_{E_1}^{CSC} = \chi_{E_1}^{CS} \cos M\alpha - \chi_{E_1}^{CC} \sin M\alpha \end{cases}$$

$$E_2 \begin{cases} \chi_{E_2}^{CCC} = \chi_{E_2}^{CC} \cos M\alpha + \chi_{E_2}^{CS} \sin M\alpha \\ \chi_{E_2}^{CSC} = \chi_{E_2}^{CS} \cos M\alpha - \chi_{E_2}^{CC} \sin M\alpha \end{cases} \tag{50}$$

$$E_3 \begin{cases} \chi_{E_3}^{CCC} = \chi_{E_3}^{CC} \cos M\alpha + \chi_{E_3}^{CS} \sin M\alpha \\ \chi_{E_3}^{CSC} = \chi_{E_3}^{CS} \cos M\alpha - \chi_{E_4}^{CC} \sin M\alpha \end{cases}$$

$$E_4 \begin{cases} \chi_{E_4}^{CCC} = \chi_{E_4}^{CC} \cos M\alpha + \chi_{E_3}^{CS} \sin M\alpha \\ \chi_{E_4}^{CSC} = \chi_{E_4}^{CS} \cos M\alpha - \chi_{E_3}^{CC} \sin M\alpha \end{cases}$$

Finally, the fourfold degenerate irreducible representation components may be written as:

$$G_1 \begin{cases} \chi_{G_1}^{CCC} = [\chi_{G_1}^{CC} + \chi_{G_3}^{CC}] \cos M\alpha \\ \chi_{G_1}^{SSC} = [\chi_{G_1}^{SS} + \chi_{G_3}^{SS}] \cos M\alpha \\ \chi_{G_1}^{CSS} = [\chi_{G_2}^{CS} - \chi_{G_4}^{SC}] \sin M\alpha \end{cases}$$

$$G_2 \begin{cases} \chi_{G_2}^{CCS} = [\chi_{G_1}^{CC} + \chi_{G_3}^{CC}] \sin M\alpha \\ \chi_{G_2}^{SSS} = [\chi_{G_1}^{SS} + \chi_{G_3}^{SS}] \sin M\alpha \\ \chi_{G_1}^{CSS} = [\chi_{G_2}^{CS} - \chi_{G_4}^{SC}] \cos M\alpha \end{cases} \tag{51}$$

$$G_3 \begin{cases} \chi_{G_3}^{CCC} = [\chi_{G_1}^{CC} + \chi_{G_3}^{CC}] \cos M\alpha \\ \chi_{G_3}^{SSC} = [\chi_{G_1}^{SS} + \chi_{G_3}^{SS}] \cos M\alpha \\ \chi_{G_3}^{CSS} = [\chi_{G_2}^{CS} - \chi_{G_4}^{SC}] \sin M\alpha \end{cases}$$

$$G_4 \begin{cases} \chi_{G_4}^{CCS} = [\chi_{G_1}^{CC} + \chi_{G_3}^{CC}] \sin M\alpha \\ \chi_{G_4}^{SSC} = [\chi_{G_1}^{SS} + \chi_{G_3}^{SS}] \sin M\alpha \\ \chi_{G_4}^{CSS} = [\chi_{G_2}^{CS} - \chi_{G_4}^{SC}] \cos M\alpha \end{cases}$$

The non-degenerate symmetry eigenvectors (49) are identical to those of non-planar pyrocatechine(40) except for the threefold multiplicity. The twofold

and fourfold degenerate eigenfunction components are linear combinations of the twofold and fourfold degenerate eigenfunction components of planar acetone.

The energy potential energy function of the Hamiltonian operator for the double C_{3v} internal rotation and wagging may be written in terms of A_1 symmetry eigenvectors [37]:

$$V(\theta_1,\theta_2,\alpha) = \sum_K \sum_L \sum_M A^{CCC}_{KLM}(\cos 3K\theta_1 \cos 3L\theta_2 + \cos 3L\theta_1 \cos 3K\theta_2)\cos M\alpha +$$
$$+ \sum_K \sum_L \sum_M A^{SSC}_{KLM}(\sin 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \sin 3K\theta_2)\cos M\alpha +$$
$$+ \sum_K \sum_L \sum_M A^{CSS}_{KLM}(\cos 3K\theta_1 \sin 3L\theta_2 - \sin 3L\theta_1 \cos 3K\theta_2)\sin M\alpha \qquad (52)$$

This expression for the potential energy is the same as that of pyrocatechin except for the threefold periodicity. In addition, it takes the same form as that of planar acetone (47) when α is zero. In the same way, it takes the form of the potential energy unction of rigid pyramidal dimethylamine when α is a constant different from zero [35,36]:

$$V(\theta_1,\theta_2,) = \sum_K \sum_L \sum_M A^{CC}_{KL}(\cos 3K\theta_1 \cos 3L\theta_2 + \cos 3L\theta_1 \cos 3K\theta_2) +$$
$$+ \sum_K \sum_L A^{SS}_{KL}(\sin 3K\theta_1 \sin 3L\theta_2 + \sin 3L\theta_1 \sin 3K\theta_2) +$$
$$+ \sum_K \sum_L A^{CS}_{KL}(\cos 3K\theta_1 \sin 3L\theta_2 - \sin 3L\theta_1 \cos 3K\theta_2)$$

It is noticeable that the NRG theory developed here furnishes two different groups for the double rotation in planar pyrocatechin (or acetone), and the double rotation and wagging mode in non-planar pyrocatechin (or pyramidal acetone). The group structures, however, are seen to be the same. The Longuet-Higgins'theory yields indeed the same Molecular Symmetry Groups for both pyramidal and planar systems. As a result, the NRG theory is seen to furnish a more detailled information about the dynamics of the non-rigid systems.

5. COMPARISON BETWEEN THE LONGUET- HIGGINS', ALTMANN'S AND NON-RIGID GROUPS

After introducing the Non-Rigid Group Theory, and the main operations used for dealing with some simple molecules, let us compare this theory with those of Longuet-Higgins [5] and Altmann [10,11]. For this purpose, let us consider the hypothetical examples of a distorted methyl halide molecule, and the orthoformic acid, advanced by Watson in an earlier critical study [38].

5.1. Distorted methyl-halide

Consider an excited electronic state of a methyl halide molecule, $CH_3 - X$, that presents three isoenergetic preferred conformations in which one of the $C - H$ bond is shorter than the other two. Due to the tunneling effect, this deformation is able to rotate from one bond to another one. as illustrated in fig. 8:

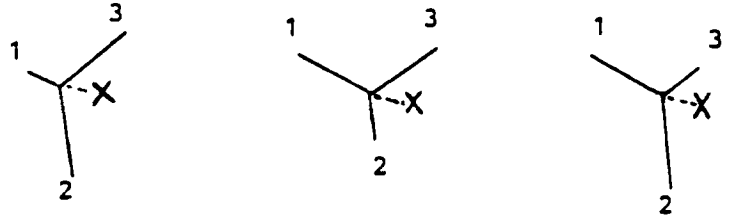


Figure 8: The three conformations of the distorted methyl halide.

The Hamiltonian operator for such a movement is easily derived, if we consider:

a) The threefold periodicity of the motions;

b) The invariance of the energy on the sense of the rotation.

As a result, a threefold potential energy function, which does not contain any antisymmetric term (sine function), may be written for the Hamiltonian operator:

$$H = -\frac{\partial}{\partial\alpha} B(\alpha) \frac{\partial}{\partial\alpha} + \sum_{K}^{\infty} A_K^C \cos 3K\alpha \tag{53}$$

where α is the rotation angle for the displacement of the deformation.

The restricted NRG corresponding to such a movement will contain all the interconversion operations that commute with the Hamiltonian operator (53),

i.e.:

$$G_{rNRG} = (C_3^I \wedge U^I) \sim C_{3v} \tag{54}$$

which is a group of order 6 isomorphic to the C_{3v} symmetry point group.

In this group, U^I and C_3^I are the simple switch and threefold subgroups (15-b) and (16-b), respectively. Notice that the C_3 operations do not commute with U.

A complete study of distorted methylhalide, should to consider the external rotation, as well as the coupling between the external and internal motions. The full Hamiltonian operator may be easily formulated, when we consider:

a) The threefold periodicity of the internal motion;

b) The invariance of the total energy under a simultaneous change of the sense of the rotation of both the deformation and the molecule.

The full NRG corresponding to such a double movement is then written as:

$$G_{fNRG} = (C_3^I \wedge V^I) \sim C_{3v} \tag{55}$$

which is again a group of order six, isomorphic to C_{3v}. In this expression the C_3^I subgroup retains the symmetry of the deformation displacement motion, and V^I, defined by (28), retains the symmetry existing in the coupling between the external and internal motions.

Let us now consider Longuet-Higgins point of view. Following the example of $CH_3 - BH_2$, the Molecular Symmetry Group is such a system and is easily written down. It will contain all the feasible operations:

$$G_6 = [(1) + (123) + (132) + (23)^* + (13)^* + (12)^*] \tag{56}$$

This group is isomorphic to the restricted and full NRG (58) and (59), as well to the C_{3v} symmetry point group.

In contrast, the Altmann's Schrödinger Supergroup is written as [39]:

$$S_6 = C_3^I \wedge C_S^G \sim C_{3v} \tag{57}$$

where C_S^G is the reflection symmetry point group:

$$C_S^G = [E + C_S]$$

C_S is the reflection operation with respect to the symmetry plane of the molecule.

Notice that all the groups considered have the same structure, but they are not identical. In particular, the restricted and full NRG are not identical between themselves, or to the Schrödinger Supergroup because $U^I \neq V^I \neq C_S^G$.

5.2. Orthoformic Acid

As a second example, let us consider the hypothetical orthoformic acid molecule $HC(OH)_3$, which does not occur naturely although esters are known. Let us suppose that the lowest configurations correspond to the two C_3 symmetric as shown in Fig. 9.

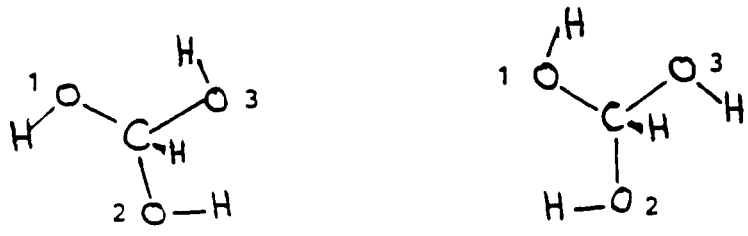


Figure 9. Orthoformic acid lowest C_3 configurations.

Due to the tunneling effect both configurations are undergoing an interconversion process. The restricted Hamiltonian operator for such a movement is easily written down if we consider:

a) A simultaneous rotation of the three hydroxyl hydrogen atoms from θ to $-\theta$ in the COH plane around the oxygen atom;

b) The invariance of the energy on the sense of the rotation.

As a result, an one-fold potential energy function, which does not contain any *sine* function, may be written for the Hamiltonian operator:

$$H = -\frac{\partial}{\partial\theta}B(\theta)\frac{\partial}{\partial\theta} + \sum_K A_K^C \cos K\theta \tag{58}$$

The restricted NRG corresponding to such a motion will contain only the identity and the simple switch operations:

$$G_{rNRG} = [E + U] = U^I \tag{59}$$

which is a group of order two.

For a complete study of symmetric orthoformic acid, we have to deduce the full NRG. The full Hamiltonian operator for the internal and external rotations, may be derived if we consider:

a) The threefold periodicity of the external rotation:

b) The invariance of the total energy under a simultaneous change of the sense of the rotation of both the rotor and the molecule.

The full NRG is then written as:

$$G_{fNRG} = C_3^I \wedge V^I \sim C_{3v} \tag{60}$$

which is again a group of order 6 isomorphic to C_{3v}. In this expression, the C_3^I subgroup retains the symmetry of the overall rotation, while the double switch subgroup, V^I, retains the symmetry existing in the coupling between the external and internal motions.

The Longuet–Higgins' Molecular Symmetry Group of this molecule is the same as that of the distorted methyl halide (56), since the number of interchangeable identical particules, i.e., the number of feasible permutations, is the same in both systems.

In contrast, the Altmann's Schrödinger Supergroup, for the symmetric orthoformic acid is expressed as [39]:

$$S_6 = U^I \times C_3^G \sim C_{3i} \tag{61}$$

which is a group of order 6 isomorphic to the C_{3i} symmetry point group, since operators C_3 and C_3^2 commute with the switch operator, U, in this particular molecular system.

5.3. Random Orthoformic Acid

In order to achieve this comparison between the Longuet-Higgins' and Altmann's theories, and the present NRG one, let us consider again the orthoformic acid molecule, but in a random configuration such as upper one shown in Fig. 10. From this configuration, it is easily seen that six different isoenergetic structures may be derived from one of them, for example, by permutations and permutations–inversions of the identical nuclei:

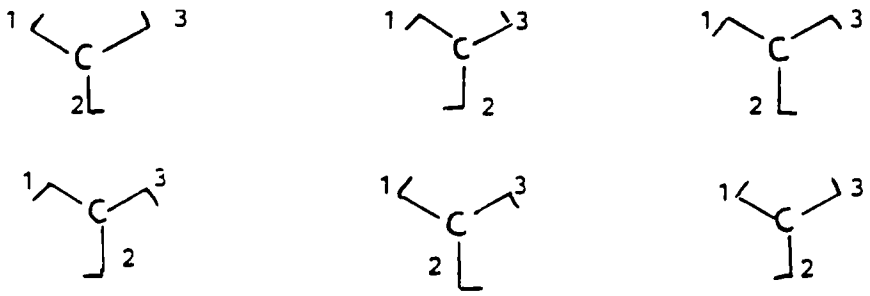


Figure 10. The six isoenergetic random orthoformic acid configurations.

These six isoenergetic configurations are not necessarily the energetically lowest ones. However, they may be assumed to be for our purposes, and if the energy barrier height between the minima are sufficiently low then interconversion movements can occur. The Hamiltonian operator corresponding to such a movement may be transcribed if we take into account:

a) The threefold periodicity of the interconversion motion along the three configurations of each file of the figure;

b) The simultaneous onefold rotation of one of hydroxylic hydrogen atoms along each column; and

c) The invariance of the energy under a simultaneous change of the sense of these motions.

As a result, the potential energy function in the Hamiltonian operator may be developed in terms of symmetric products of trigonometric functions, one of periodicity three and of periodicity one, respectively:

$$H = T(\theta, \alpha) + \sum_K \sum_L [A_{KL}^{CC} \cos 3K\theta \cos L\alpha + A_{KL}^{SS} \sin 3K\theta \sin L\alpha] \qquad (62)$$

where θ and α are the rotation angles for the displacement of the distortion, and the simultaneous rotations of the hydroxyl hydrogen atoms, respectively.

The complete set of interconversion operations, which commute with this Hamiltonian operator, is given by the semi-direct product:

$$G_{rNRG} = C_3^I \wedge V^I \sim C_{3v} \qquad (63)$$

which is a group of order 6 isomorphic to C_{3v}.

The full NRG for this molecule is identical to the restricted one, because the random orthoformic acid does not possess any symmetry.

The Longuet–Higgins' Molecular Symmetry Group in this third example is the same as the previous one (56), because the number of feasible permutations of identical particles is the same.

The Altmann' Schrödinger Supergroup takes now the form [39]:

$$S_6 = (C_3^I \wedge U^I) \sim C_{3v} \qquad (64)$$

because the Altmann 's theory does not taken into account the external motions.

5.4. Discussion

Since the NRG theory uses the same formalism as the Altmann's for describing the different operations, and the restricted NRG is limited to the

internal motion, we could expect that the restricted NRG will coincide with the Altmann's Isodynamic group (see Section 2.2.). This is not necessarily true. An example is the restricted NRG for the distorted methylhalide (54) which appears to be different from the Altmann's Isodynamic group (57).

Similarly, because the full NRG considers the external and internal motions as well as the coupling between them, the full NRG does not coincide with the Altmann' Schrödinger Supergroup, even when the molecule under study has no symmetry such as in the case of a random orthoformic acid.

Notice that the full NRG's of this series of molecular systems may not be written as a semi-direct product of a restricted NRG for the internal motions and a symmetry point subgroup for the external rotation, as Altmann presented [10], because of the coupling between both motions. The full NRG's, however, possess the same group structure as the Longuet-Higgins' Molecular Symmetry Group, in all the cases.

If we consider now the example of the random orthoformic acid, it is evident that the Longuet-Higgins, Altmann and full NRG theories yield groups with the same structure, although the groups are different. On behalf of the Schönfliess formalism, however, the full NRG theory is able to take into account explicitly the coupling between the external and internal rotations. This example illustrates clearly the power of the full non-rigid group theory.

The results are sumarized in Table 9.

Table 9: Groups and group structures in the Altmann's, restricted and full NRG theories, and group structure in the Longuet-Higgins scheme for distorted methyl halide (DMH), symmetric orthoformic acid (SOA), and random orthoformic acid (ROA).

Molecules	Altmann	restricted NRG	full NRG	L-H
DMH	$C_3^I \wedge C_s^G \sim C_{3v}$	$C_3^I \wedge U^I \sim C_{3v}$	$C_3^I \wedge V^I \sim C_{3v}$	$G_6 \sim C_{3v}$
SOA	$U^I \wedge C_3^G \sim C_{3i}$	$U^I \sim C_s$	$C_3^I \wedge V^I \sim C_{3v}$	$G_6 \sim C_{3v}$
ROA	$C_3^I \wedge U^I \sim C_{3v}$	$C_3^I \wedge V^I \sim C_{3v}$	$C_3 \wedge V^I \sim C_{3v}$	$G_6 \sim C_{3v}$

6. FURTHER COMPARISONS

In sections 3 and 4, we have considered non-rigid molecules with a solid reference frame. In order to achieve the comparison between the NRG's and the Longuet–Higgins and Altmann's groups, let us consider some very symmetric systems, such as linear molecules with equivalent rotors, in which the reference frame is only a rotational axis, or centro–symmetric molecules with permutational rearrangements around a central point, in which the reference frame is a single atom.

6.1. Linear Symmetric Molecules

In sections 2 and 5., we have already considered linear molecules: boron difluoride methyl and distorded methylhalide. Let us consider next a more symmetric linear molecule bearing two equivalent rotors: the ethane molecule.

6.1.1 Ethane

The restricted Hamiltonian operator for the internal rotation in ethane, in a random conformation, may be easily written by taking into account:

a) The threefold periodicity of the internal rotation of one of the methyl group with respect to the other;
b) The invariance of the energy on the sense of the rotation.
As a result we have:

$$H = -\frac{\partial}{\partial\theta}B(\theta)\frac{\partial}{\partial\theta} + \sum_K A_K^C \cos 3K\theta \tag{65}$$

where θ is the rotation angle.

As a result, the restricted NRG of the operations which commute with this Hamiltonian operator is:

$$G_{rNRG} = (C_3^I \wedge U^I) \sim C_{3v} \tag{66}$$

which forms a group of order 6 isomorphic to C_{3v}.

In order to derive the full NRG of ethane, let us consider this molecule in a laboratory axis system, fixed in space, with the origin located in the center of mass and the Z axis along its C_3 axis. Let us remark that, in these conditions, the dynamical symmetry properties of ethane may be deduced as follows [22]:

a) The invariance of the energy with respect to a C_3 rotation of whichever of the rotors;

b) Both methyl group are equivalent;

c) The invariance of the energy under a simultaneous inversion of the rotation sense of both rotors.

Comparing these properties with those of acetone (4.1.), the full NRG may be expressed as:

$$G_{fNRG} = (C_3^I \times C_{3'}^I) \wedge [W^I \times V^I] \sim G_{36} \tag{67}$$

which is a group of order 36 formally identical to that of acetone (42). The invariance conditions are indeed the same.

In this expression, the C_3^I and $C_{3'}^I$ subgroups correspond to the threefold rotations of each methyl groups, with respect to a laboratory reference axis. The W^I exchange subgroup corresponds to the equivalence of the rotors, and the V^I subgroup to the invariance of the energy with respect a change of the rotation sense.

The Longuet-Higgins' Molecular Symmetry Group for ethane is given in ref. [5]. It is a group of order 36, which may be written as a direct product of two subgroups of order 6. If the hydrogen atoms of each methyl radical are labelled from 1 to 3, and from 4 to 6, the elements of these subgroups are [19]:

$$\begin{array}{llll} \text{a) } (1) & (123)(465) & (14)(25)(36) \\ & (132)(456) & (15)(26)(34) \\ & & (16)(24)(35) \\ \text{b) } (1) & (123)(456) & (14)(26)(35)^* \\ & (132)(365) & (15)(24)(36)^* \\ & & (16)(25)(34)^* \end{array} \tag{68}$$

Each of these subgroups are isomorphic to C_{3v}. The Molecular Smmetry Group for ethane is identical to that of acetone, because the number of feasible permutations and permutations-inversions of identical particles is the same in both systems, and isomorphic to the full NRG Groups.

The Altmann's Schrödinger Supergroup for ethane is provided in ref. [19]. Considering that ethane in a random conformation possesses still one C_3 and three C_2 symmetry axes, see Fig. 11, the Schrödinger Group of this molecule is:

$$D_3 = [E + 2C_3 + 3C_2] = [C_3^G \times C_2^G] \tag{69}$$

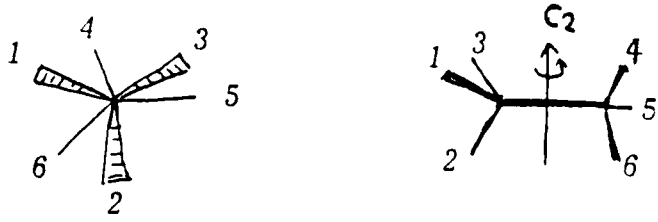


Fig. 11.- The ethane molecule in stereographic projection and presented perpendicularly to the C-C bond. Labbeling of the hydrogen atoms and symmetry axes.

On the other hand, according to the $CH_3 - BF_2$ example, the isodynamic group will contain the C_3^I subgroup (16-b) and the simple switch subgroup (15-b). As a result, the Altmann isodynamic group will be written as:

$$I = [C_3^I \wedge U^I] \sim C_{3v} \tag{70}$$

which is a group of order 6, isomorphic to C_{3v} and identical to the restricted NRG.

The Schrödinger Supergroup then has the form:

$$S_{36} = (C_3^I \wedge U^I) \wedge D_3 \sim G_{36} \tag{71}$$

which is a group of order 36 isomorphic with the full NRG and the Longuet-Higgins'one, G_{36}.

In the present case of ethane, it is evident that the Longuet-Higgins. Altmann's and NRG theories yield groups with the same group structure. The groups, however, are not the same in the Altmann's and NRG approach. Let us remark that the V^I and W^I subgroups in the fNRG describe the external and internal rotations with the interaction between them, while the U^I subgroup in the Schrödinger Supergroup only considers the internal rotation.

Notice that because of the existence of the W^I subgroup,i.e., the equivalence between the rotors, in a linear molecule such as ethane, the rotating moiety may be whichever of the methyl groups, which may rotate from 0° to 360°. As a result, the external and internal rotation solutions have to be classified according to a two-valued G_{36} group [8,9].

6.2. Centro–Symmetric Molecular Systems

Centro–symmetric molecular systems, such as PF_5 , XeF_6, transition metal coordination complexes, and polytopal molecules, in which the internal conversion motions can be described as permutational rearrangements around a single atom, furnish another example where the Longuet-Higgins', NRG's and Altmann's theories differ. The Schrödinger Supergroup of symmetric PF_5 was seen, indeed, to be not factorizable into two subgroups as expected in the Altmann's theory [20,25]. This kind of non-rigid molecular systems were extensively studied into the Longuet-Higgins' formalism [40-45].

6.2.1. Phosphorus pentafluoride

In order to compare the theories, let us consider the example of symmetric phosphorus pentafluoride, which exhibits a pyramidal structure with triangular basis. In this structure, the identical fluorine nuclei do not have the same locations. Two kinds of positions exist: the tops of the bipyramid, i.e., the axial positions, and the tops of the equilateral triangular basis, i.e., the equatorial positions. These positions are shown in Fig. 12.

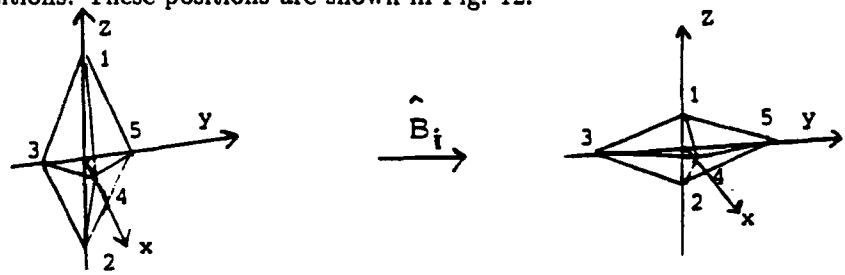


Figure 12.- Phosphorus pentafluoride structure and interconversion mechanism of Berry [46].

A few years ago Berry suggested that phosphorus pentafluoride can undergo an interconversion process in which two axial fluorine atoms are transformed into two equatorial ones, and reciprocally [46]. This mechanism is illustrated in Fig. 12. There are various experimental data supporting with such a mechanism, although other mechanisms were proposed.

In fact, it is easily seen that there exist 20 isoenergetic configurations, and 20 interconversion operations connecting them. The Altmann's isodynamic group could be written then in terms of 20 double rotation operations according to [46]:

$$I_{20} = B^I \tag{72}$$

Taking into account the symmetry point group of the molecule, D_{3h}, the Schrödinger Supergroup is expressible as:

$$S_{240} = B^I \wedge D_{3h} \tag{73}$$

which should be a group of order 240.

If the structure of the possible isodynamic group (72) is analyzed, it is found that the set of the double rotations is not closed, i.e., some products give rise to overall rotations of the molecule as a whole.

This result was foreseen earlier by using the Longuet-Higgins formalism. The Molecular Symmetry Group for PF_5 may be indeed written as the permutation group of five identical nuclei multiplied by the inversion subgroup [20]:

$$G_{240} = S_5 \times [E + E^*] \tag{74}$$

which is a group of order 240 in accordance with (73). These 240 permutations and permutations-inversions are known as the Berry operations. Let us now remark that the symmetric group, S_5, possesses one and only one unvariant subgroup, i.e., the Alternate group of the even permutations, A_5, of order 60. As a result G_{240} of (74) cannot be factorized into two subgroups of order 20 and 12 respectively, as presented in expression (73).

The symmetric PF_5 molecule is a clear example where the Schrödinger Supergroup cannot be factorized into two subgroups, as expected by the Altmann's theory. In this case, the internal rotations around a single atoms induce overall rotations of the molecule as a whole, i.e. the internal and external motions are not separable.

In the case of centro–symmetric molecules like phosphorus pentafluoride, the restricted NRG cannot be written out, furthermore the full NRG is not easy to determine. This should be expressed in terms of double rotation operations, B_i, and two- and threefold overall rotations.

7. LOCAL NON–RIGID GROUPS: THE LOCAL HAMILTONIAN OPERATOR

Up to now we have rigorously considered the restricted Hamiltonian operator for some molecular motions, in which all the intramolecular interactions were taken into account. In the following, we shall neglect some of these interactions we suppose small enough, and we shall deal with approximate or *local*

Hamiltonian operators. We shall see that with such approximate operators, we shall be able to deduce new and larger groups, which shall permit to introduce additional simplifications into the Hamiltonian matrix solutions. These new groups will be called **Local Restricted Non–Rigid Groups** [21,22]. The idea of using approximate Hamiltonian operators was already for warded by Bunker [8].

7.1. Benzaldehyde without the cog–wheel effect

Our first simple example of a local rNRG will consider the benzaldehyde molecule, where the two moving parts are assumed to move independently of the torsion sense of each other, i.e., we assure that there is no cog–wheel effect between both movements. This effect is known indeed to be very small.

The restricted Hamiltonian operator corresponding to such a moving system, takes the simpler form:

$$H^L = [-B_\theta \frac{\partial^2}{\partial\theta^2} - B_\alpha \frac{\partial^2}{\partial\theta^2}] + \sum_K \sum_L A_{KL}^{CC} \cos 2K\theta \cos 2L\alpha \qquad (75)$$

If we compare this local Hamiltonian operator with the exact one (25), we can verify that the kinetic interaction terms, as well as the potential ones in *sine* × *sine* are dropped. The second order sine terms are known to introduce the cogwheel effect in the potential energy function, because the *sine* function depends on the rotation sense [34,35].

Since the two motions do not depend anymore on the rotation sense of each other, the double switch operator of (26) may be written as a product of two simple switch operators (15-a) acting independently on each single motion. As a result the switch subgroup takes the particular form:

$$(U_1^I \times U_2^I) = [E + U_1 + U_2 + U_1U_2] \qquad (76)$$

where

$$U_i f(\theta_i) = f(-\theta_i) \quad and \quad V = U_1U_2. \qquad (77)$$

The restricted local NRG may be thus written as

$$G_{rNRG}^L = [C_2^I \times (U_1^I \times U_2^I)] \sim C_{2v} \times C_S \qquad (78)$$

which is a group of order 8, isomorphic to the point group product $C_{2v} \times C_S$.

The character table for (78) is obtained by multiplying the character table of the C_{2v} group by that of the C_S, taking into account that C_{rNRG}^L is now an Abelian group.

From the character table, the symmetry eigenvectors for the double torsion without cog-wheel effect in benzaldehyde are easily derived on the basis of the double free rotor solutions. It is found:

$$\begin{array}{ll} \chi_{A_1} = \cos 2K\theta \cos L\alpha & \chi_{B_1} = \cos(2K+1)\theta \cos L\alpha \\ \chi_{A_1'} = \sin 2K\theta \sin L\alpha & \chi_{B_1'} = \sin(2K+1)\theta \sin L\alpha \\ \chi_{A_2} = \cos 2K\theta \sin L\alpha & \chi_{B_2} = \cos(2K+1)\theta \sin L\alpha \\ \chi_{A_2'} = \sin 2K\theta \cos L\alpha & \chi_{B_2'} = \sin(2K+1)\theta \cos L\alpha \end{array} \tag{79}$$

The same symmetry eigenvectors as in the case of benzaldehyde with cog-wheel effect are encountered (29), but in the present case their group properties are different.

Using these symmetry eigenvectors the Hamiltonian matrix is factorized in eight boxes instead of four. This feature introduces a new simplification in the solution of the double rotor Schrödinger equation, paying a very small loss of accuracy.

7.2.Pyrocatechin without a cog–wheel effect

A local Hamiltonian may be presented for the pyrocatechin case. It just a matter of a mathematic model, since the interaction between the two hydroxyl groups ought to be relatively strong, due to an intramolecular hydrogen bond formation.

The restricted local Hamiltonian operator should be written as [21]:

$$H^L = [-B_1 \frac{\partial^2}{\partial \theta_1^2} - B_2 \frac{\partial^2}{\partial_2^2}] + \\ + \sum_K \sum_L A_{KL}^{CC} [\cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2] \tag{80}$$

where the kinetic interaction terms and *sine* × *sine* terms are dropped.

Following the example of benzaldehyde the local rNRG is expressed as:

$$G_{rNRG}^L = (C_I^I \times C_I'^I) \wedge [W^I \wedge (U_1^I \times U_2^I)] \sim C_{4v} \tag{81}$$

which is a group of order eight isomorphic to the symmetry point group C_{4v}. W^I is defined by (32) and $(U_1^I \times U_2^I)$ by (76). Both subgroups do not commute because:

$$WU_1 = U_2 W \tag{82}$$

It is easily seen that U_1 and U_2 form a class, and W and $WV = WU_1U_2$ another one. Finally, V which commutes with all the group elements forms by itself a class. As a result, the rNRG (81) is seen to be isomorphic with C_{4v}, and the character table may be written as presented in Table 10.

Table 10: Character table for the double rotation without the cog–wheel effect in pyrocathechin.

	E	V	$2U_i$	W, WV	$2WU_i$
A_1	1	1	1	1	1
A_2	1	1	-1	1	-1
B_1	1	1	1	-1	-1
B_2	1	1	-1	-1	1
E	2	-2	0	0	0

From this character table, the symmetry eigenvectors may be deduced for the double internal rotation without cog–wheel effect in pyrocatechin. It is found:

$$\begin{aligned}
\chi_{A_1} &= \cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2 \\
\chi_{A_2} &= \sin K\theta_1 \sin L\theta_2 + \sin L\theta_1 \sin K\theta_2 \\
\chi_{B_1} &= \cos K\theta_1 \cos L\theta_2 - \cos L\theta_1 \cos K\theta_2 \\
\chi_{B_2} &= \sin K\theta_1 \sin L\theta_2 + \sin L\theta_1 \sin K\theta_2 \\
\chi_E &= \begin{cases} \cos K\theta_1 \sin L\theta_2 + \sin L\theta_1 \cos K\theta_2 \\ \cos K\theta_1 \sin L\theta_2 - \sin L\theta_1 \cos K\theta_2 \end{cases}
\end{aligned} \tag{83}$$

They are the same symmetry eigenvectors as those of pyrocathechin with the cog–wheel effect (34), but the group properties are different.

7.3. Non–planar Pyrocatechin Without the Cog–wheel Effect

Finally, let us consider a three–dimension problem, i.e., the double rotation and wagging in pyrocatechin. Three different possibilities may be considered:

a) There is no cog–wheel effect between the rotors and the wagging atom;

b) There is no cog–wheel effect between the rotors themselves;

c) There is no cog–wheel effect between any of the moving parts.

7.3.1. Non planar Pyrocatechin without the cog–wheel effect between the rotors and the wagging atoms

When there is no cog–wheel effect between the rotors and the wagging atoms the cross–products between the cosine and sine functions, which appear in the potential (41) of the complete Hamiltonian operator may be dropped. So we have:

$$H^L = [-B_1\frac{\partial^2}{\partial\theta_1^2} - 2B_{12}\frac{\partial}{\partial\theta_1\partial\theta_2} - B_1\frac{\partial^2}{\partial\theta_2'} - B_3\frac{\partial^2}{\partial\alpha^2}]+$$
$$+\sum_K\sum_L\sum_M A^{CCC}_{KLM}[\cos K\theta_1\cos L\theta_2 + \cos L\theta_1\cos K\theta_2]\cos M\alpha \qquad (84)$$
$$+\sum_K\sum_L\sum_M A^{SSL}_{KLM}[\sin K\theta_1\sin L\theta_2 + \sin L\theta_1\sin K\theta_2]\cos M\alpha$$

As a result the local Hamiltonian operator does not commute solely with the triple switch operator (35), but also with the double switch, V, and simple switch, U, independently. So, the local group may be easily deduced replacing in (39) the triple switch subgroup $(VU)^I = [E+VU]$ by the subgroup product $V^I \times U^I = [E+V]\times[E+U]$. After some straightforward simplifications we obtain:

$$G^L_{rNRG} = [C_1^I \times C_{1'}^I] \wedge [W^I \times V^I \times U^I] \qquad (85)$$

which is a group of dimension eight isomorphic to that of benzaldehyde without the cog–wheel effect (78). The operation U commutes with all the elements of the group.

The symmetry eigenvectors of such a group are easily deduced multiplying those of planar pyrocatechin (34) by $\cos M\alpha$ or $\sin M\alpha$.

7.3.2. Non-planar Pyrocatechin without the cog–wheel effect between the Rotors themselves

When there is no cog–wheel effect between the rotors themselves, but with the wagging atoms, the *sine* − *sine* symmetric binomial forms, which appear in the complete Hamiltonian operator (41), could be expected to be dropped. However, *sine* × *sine* products of the rotation angles and the wagging angle could remain:

$$H^L = \{[-B_1\frac{\partial^2}{\partial\theta_1^2} - B_1\frac{\partial^2}{\partial\theta_2^2} - B_{13}(\frac{\partial^2}{\partial\theta_1\partial_\alpha} + \frac{\partial^2}{\partial\theta_2\partial_\alpha}) - B_3\frac{\partial^2}{\partial\alpha^2}]\}+$$
$$+\sum_K\sum_L\sum_M A^{CCC}_{KLM}[\cos K\theta_1\cos L\theta_2 + \cos L\theta_1\cos K\theta_2]\cos M\alpha+ \qquad (86)$$

$$+\sum_K \sum_L \sum_M A_{KLM}^{SSS}[\sin K\theta_1 \sin L\theta_2 - \sin L\theta_1 \sin K\theta_2] \sin M\alpha$$

In such a case, the Hamiltonian operator no longer commutes with the triple switch operator (35), but only with those of the double switch ones, UU_1 and UU_2, independently. These operators are defined as products of simple switch operator acting on α and θ_1, and α and θ_2, respectively.

For constructing the local group, let us rewrite the complete non-rigid subgroup (38) as:

$$G_{rNRG} = [E + UV + WU + WV] = [(WU)^I \times (UV)^I] \tag{87}$$

and replace in (87) the triple switch operator $(UV)^I$) by the following product of subgroups:

$$(UU_1)^I \times (UU_2)^I = [E + UU_1] \times [E + UU_2]$$

We obtain:

$$G_{rNRG}^L = [C_I^I \times C_{I'}^I] \times (WU)^I \wedge [(UU_1)^I \times (UU_2)^I] \tag{88}$$

which is another group of order eight isomorphic to that of planar pyrocatechine without the cog-wheel effect (81) and the symmetry point group C_{4v}.

From the multiplication properties of the operations the character table for a non-planar pyrocatechine without the cog-wheel effects between the rotors themselves but with the wagging atoms, is easily deduced. This is given in table 11.

Table 11: Character table for the double rotation without the cog-wheel effect and wagging vibration mode in pyrocathechin.

	E	V	$2U_iU$	WU WVU	$2WU_i$
A_1	1	1	1	1	1
A_2	1	1	-1	1	-1
B_1	1	1	1	-1	-1
B_2	1	1	-1	-1	1
E	2	-2	0	0	0

From this character table the symmetry eigenvectors are easily derivable on the basis of the triple free rotor solution. It is found:

$$\chi_{A_1} = \begin{cases} (\cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2) \cos M\alpha \\ (\sin K\theta_1 \sin L\theta_2 - \sin L\theta_1 \sin K\theta_2) \sin M\alpha \end{cases}$$

$$\chi_{A_2} = \begin{cases} (\cos K\theta_1 \cos L\theta_2 - \cos L\theta_1 \cos k\theta_2) \sin M\alpha \\ (\sin K\theta_1 \sin L\theta_2 + \sin L\theta_1 \sin K\theta_2) \cos M\alpha \end{cases}$$

$$\chi_{B_1} = \begin{cases} (\cos K\theta_1 \cos L\theta_2 - \cos L\theta_1 \cos K\theta_2) \cos M\alpha \\ (\sin K\theta_1 \sin L\theta_2 + \sin L\theta_1 \sin K\theta_2) \sin M\alpha \end{cases} \tag{89}$$

$$\chi_{B_2} = \begin{cases} (\cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos k\theta_2) \sin M\alpha \\ (\sin K\theta_1 \sin L\theta_2 - \sin L\theta_1 \sin K\theta_2) \cos M\alpha \end{cases}$$

$$\chi_E = \begin{cases} (\cos K\theta_1 \sin L\theta_2 + \sin L\theta_1 \cos K\theta_2) \cos M\alpha \\ (\cos K\theta_1 \sin L\theta_2 - \sin L\theta_1 \cos K\theta_2) \sin M\alpha \\ \\ (\cos K\theta_1 \sin L\theta_2 - \sin L\theta_1 \cos K\theta_2) \cos M\alpha \\ (\cos K\theta_1 \sin L\theta_2 + \sin L\theta_1 \cos K\theta_2) \sin M\alpha \end{cases}$$

As expected, the completely symmetric eigenvectors A_1 contains the product of the sine between the rotation angles and the wagging angle, the *sine* × *sine* binomial forms being antisymmetric.

7.3.3. Non-planar pyrocatechin without any cog–wheel effect

When there are no cog-wheel effect between the rotating parts, the *sine* × *sine* products may be dropped in (41), (84) or (86). So, the local Hamiltonian operator may be expressed as:

$$H^L = \{[-B_1 \frac{\partial^2}{\partial\theta_1^2} - B_1 \frac{\partial^2}{\partial\theta_2^2} - B_3 \frac{\partial^2}{\partial\alpha^2} +$$

$$+ \sum_K \sum_L \sum_M A_{KLM}^{CCC} [\cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2] \cos M\alpha \tag{90}$$

In this special case, the Hamiltonian operator does not commute only with the triple switch, but also with the double and simple ones. As a result, the local group may be written replacing in (39) the triple switch subgroup by the triple product of the subgroup U^I, U_1^I and U_2^I.

$$U^I \times U_1^I \times U_2^I = [E + U] \times [E + U_1] \times [E + U_2] \tag{91}$$

We obtain:

$$G_{rNRG}^L = [C_i^I \times C_{i'}^I] \wedge [W^I \wedge (U^I \times U_1^I \times U_2^I)] \tag{92}$$

which is a group of dimension 16 isomorphic to the point group product $C_{4v} \times C_S$.

The symmetry eigenvectors may be deduced from those of the preceding group (89) by taking into account that the binomial products of $\cos M\alpha$ and $\sin M\alpha$ belong now to different irreducible representations.

The symmetry eigenvectors of any local NRG may be advantageously used as basis functions in a pre-diagonalization of the complete Hamiltonian matrix. For this purpose, the local NRG has only to be a larger group than the complete NRG. In the present case we have for example:

$$C_{2v} \subset C_{4v} \subset C_{4v} \times C_S \tag{93}$$

for pyrocatechin without any cog–wheel–effect, without the cog–wheel effect between the rotors themselves and with all the cog–wheel effects, respectively.

8. LOCAL NON–RIGID GROUPS FOR C_{3v} ROTOR MOLECULES

In the previous section, we considered the relatively simple non–rigid systems and we have deduce their local rNRG's for different stages of simplification of the Hamiltonian operator. Next, we shall consider molecular systems bearing a C_{3v} rotor. The study of these more complex molecules better illustrated the advantage of the using the local NRG.

8.1. Planar Acetone Without the Cog–wheel Effect

The restricted local Hamiltonian operator for the double internal rotation in acetone neglecting the cog–wheel effect between the rotors is similar to that of pyrocatechine (80), except for the periodicity of the rotor. The restricted local NRG is deduced directly from that of pyrocatechin (81), taking into account (42) [21-22]:

$$G^L_{rNRG} = [C^I_3 \times C^I_{3'}] \wedge [W^I \wedge (U^I_1 \times U^I_2)] \tag{95}$$

which is a group of order 72. Its character table may be derived easily from acetone taking into account the splitting of the V^I subgroup. This is given in Table 12.

Table 12: Character Table for the double internal rotation in acetone without cog-wheel effect.

	E	$4C_3$	$4\ C_3C_{3'}$	$6\ W$	$12\ W\ C_3C_{3'}$	$9\ V$	$6U_i$	$12\ U_iC_3$	$18WU_i$
A_1	1	1	1	1	1	1	1	1	1
A_2	1	1	1	-1	-1	1	1	1	-1
A_3	1	1	1	1	1	1	-1	-1	-1
A_4	1	1	1	-1	-1	1	-1	-1	1
E	2	2	2	0	0	-2	0	0	0
G_1	4	1	-2	0	0	0	2	-1	0
G_2	4	1	-2	0	0	0	-2	1	0
G_3	4	-2	1	2	-1	0	0	0	0
G_4	4	-2	1	-2	1	0	0	0	0

In this table, there exist four non-degenerate irreducible representations, one twofold degenerate and four fourfold degenerate. From this table and that of acetone with the cog-wheel effect, the equivalence relations between the irreducible representations of both groups are derivable. These are given in Table 13.

Table 13: Equivalence relations between the irreducible representations for the double rotation in acetone with and without cog-wheel effect.

With cog-wheel effect		Without cog-wheel effect
A_1	$\in$	$A_1 + A_1'$
$A_2 + A_3$	$=$	E
A_4	$\in$	$A_2 + A_2'$
$E_1 + E_3$	$=$	G_3
$E_2 + E_4$	$=$	G_4
G	$\in$	$G_1 + G_2$

The potential energy function for acetone without cog-wheel effect between the rotors is then written in terms of symmetry eigenvectors corresponding to the completely symmetric representation:

$$V(\theta_1, \theta_2) = \sum_K \sum_L A_{KL}^{CC}(\cos 3K\theta_1 \cos 3L\theta_2 + \cos 3L\theta_1 \cos 3K\theta_2) \qquad (96)$$

expression which does not contain any *sine* × *sine* products.

Using the symmetry eigenvectors the Hamiltonian matrix is factorable into 32 boxes instead of 16, with a very slight loss of accuracy.

8.2. Pyramidal Acetone-like Molecules Without the Cog–Wheel Effect

The pyramidal acetone-like molecules, such as acetone in an electronic excited state, dimethylamine, yield good examples of applications of the local groups to C_{3v} rotor molecules. As in non–planar pyrocatechin, three different local Hamiltonian operators may be considered.

8.2.1. Non planar acetone-like molecules without the cog–wheel effect between the rotors and the wagging atom

Acetone in an electronic excited state, in which the oxygen atom is wagging out-of-plane, furnishes an example of such a case. The corresponding Hamiltonian operator is given by expression (84) in which the threefold periodicity of the rotors are introduced. Similarly, the local rNRG may deduced from the expression (85):

$$G_{rNRG}^{L} = [C_3^I \times C_{3'}^I] \wedge [W^I \times V^I \times U^I] \qquad (97)$$

which is a group of order 72 isomorphic to the group product $G_{36} \times C_2$ of p-xylene [35,36], since U commutes will all the elements of the group.

The eigenvectors of such a group may be derived easily from those of planar acetone (96) multiplying them by $\cos M\alpha$ or $\sin M\alpha$.

8.2.2. Non–planar acetone-like molecules without the cog–wheel effect between the rotors themselves

Dimethylamine in which the methyl groups are rotating in a first approach without a cog–wheel effect between them. The wagging of the N-hydrogen atom, on the contrary, induces an inversion of the whole molecule, that cannot be easily neglected.

The corresponding Hamiltonian operator is given essentially by expression (86), and its local rNRG by expression (88), where the threefold periodicity of the rotors is introduce: so, we have for the local rNRG:

$$G'^{L}_{rNRG} = [C_3^I \times C_{3'}^I] \wedge \{(WU)^I \wedge [(UU_1) \times (UU_2)^I]\} \tag{98}$$

which is another group of order 72 isomorphic to that of planar acetone without a cog-wheel effect (95).

The symmetry eigenvectors may be constructed from those of the former group, by multiplying by $\cos M\alpha$ or $\sin M\alpha$, and gathering the products by pairs according to their symmetry properties with respect to operations of the subgroup $\{(WU)^I \wedge [(UU_1)^I \times (UU_2)^I]\}$ of (98), as in the case of non-planar pyrocatechine (89).

8.2.3. Non planar acetone–like molecules without any cog-wheel effect

The local Hamiltonian operator and the local rNRG are given essentially by expressions (90) and (92), respectively, in which the threefold periodicity of the rotors has to be introduced. So, the local rNRG has the simple form:

$$G''^{L}_{rNRG} = [C_3^I \times C_{3'}^I] \wedge [W^I \wedge (U^I \times U_1^I \times U_2^I)] \tag{99}$$

which is a group of order 144 isomorphic to the group product G_{72} of (95) $\times C_2$.

As in the previous case of pyrocatechin, the symmetry eigenvectors of such a group may be deduced from those of the preceding group (98) by taking into account in the present case the products by $\cos M\alpha$ and $\sin M\alpha$ belonging to different irreducible representations.

9. RESTRICTED AND LOCAL NON–RIGID GROUPS

The restricted Hamiltonian operator for a non–rigid molecular system may be regarded as a special case of a local Hamiltonian operator in which the external rotation term has been dropped. This case holds only when the external and internal motions are separable. Let us consider some typical examples.

9.1. Non–rigid Systems without Symmetry in a Random Configuration

When the molecule, in a random configuration, does not have any symmetry, its restricted NRG is expected to be isomorphic to the full NRG (see section 2.4.).

9.1.1. Phenol

Let us return to the example of phenol (Fig.3) in which the hydroxyl moiety is able to rotate around the C_2 axis of the phenyl ring. This molecular system has no symmetry in a random configuration. The full Hamiltonian operator of such a system is written as (18):

$$H(X_e, X_i) = T(X_e) + T(X_e, X_i) + H(X_i) \tag{100}$$

where $H(X_i)$ is the restricted Hamiltonian operator (21).

In order to deduce its full NRG let us consider:

a) The twofold periodicity of the rotation of the hydroxylic group;

b) The invariance of the total energy under a simultaneous change of the sense of the rotations of both the rotor and the molecule.

The full NRG corresponding to such a double movement is then written as:

$$G_{fNRG} = (C_2^I \wedge V^I) \sim C_{2v} \tag{101}$$

where V^I is the double switch subgroup (27) corresponding to the external and internal rotations.

Let us now drop the kinetic interaction operator in (100). The full Hamiltonian has now the approximate form:

$$H^L(X_e, X_i) = T(X_e) + H(X_i) \tag{102}$$

in which the external and internal variables are separable.

According to (76), the *local* full NRG may be written as:

$$G^L_{fNRG} = [C_2^I \wedge (U_e^I \times U_i^I)] \sim C_{2v} \times C_S \tag{103}$$

In this expression, U_e^I and U_i^I are the single switch subgroups (15-b) corresponding to the external and internal rotation angles, respectively. This group is larger than the *exact* full NRG (101), i.e., a group of order 8 isomorphic to $C_{2v} \times C_S$.

Taking into account that there are no interaction between the external and internal rotation, the switch subgroup corresponding to external rotation may be completely factorized, so that:

$$G^L_{fNRG} = [C_2^I \wedge U_i^I] \times U_e^I \tag{104-a}$$

The *local* full NRG appears then as a direct product of two subgroups: the restricted NRG corresponding to the internal motions, and a switch subgroup, U_e^I, corresponding to the external rotation, in accordance with equation (20):

$$G^L_{fNRG} \sim [C_2^I \wedge U_i^I] \times C_S \tag{104-b}$$

As to be expected, the restricted NRG is seen to be isomorphic to the full NRG, and the external symmetry rotation group is found to be isomorphic to the symmetry point group of the molecule, C_S, in its most symmetric configuration. Similar expressions may be deduced for benzaldehyde, pyrocatechin and acetone.

9.2. Non–rigid molecules with symmetry in a random configuration

When the non–rigid molecule in a random configuration retains some symmetry, the restricted NRG is expected to be smaller than the full NRG.

9.2.1. Distorted methylhalide

Let us return to the example of the distorted methylhalide (Fig.8) in which a C_S symmetry plane is retained. The full Hamiltonian operator of such a system will be written as (100). As found in Section 5, the full NRG of such a molecular system is written as (55):

$$G_{fNRG} = (C_3^I \wedge V^I) \sim C_{3v} \tag{105}$$

which is group of order 6 isomorphic with the symmetry point group C_{3v}. Neglecting the kinetic interaction between the external and internal rotations in the full Hamiltonian operator, the *local* full NRG may be expressed taking into account (76) as:

$$G_{fNRG}^L = [C_3^I \wedge (U_e^I \times U_i^I)] \sim C_{3v} \times C_S \tag{106}$$

This *local* group is larger than the *exact* full NRG (101), i.e., a group of order 12 isomorphic to $C_{3v} \times C_S$.

$$G_{fNRG}^L = [C_3^I \wedge U_i^I] \times U_e^I \tag{107-a}$$

The *local* full NRG appears then as a direct product of two subgroups: the restricted NRG corresponding to the internal motions, and a single switch subgroup, U_e^I, which corresponds to the external rotation. The restricted NRG is found to be smaller than the full NRG and the external rotation one isomorpic to the symmetry point group of the molecule, C_S, in its most symmetric configuration. As a result, it may be written as in the case of phenol:

$$G_{fNRG}^L \sim [C_3^I \wedge U_i^I] \times C_S \tag{107-b}$$

in accordance with expression (20).

9.2.2. Symmetric orthoformic acid

Let us now consider again the symmetric orthoformic acid (Fig.9), in which a C_3 symmetry axis is retained. The full Hamiltonian operator of such a molecule has the same general form (100). As found in section 5, the full NRG of symmetric orthoformic acid is written as (60):

$$G_{fNRG} = [C_3^I \wedge V^I] \sim C_{3v} \tag{108}$$

Neglecting the interactions between the external an internal rotations, the *local* full NRG may be written taking into account (76) as:

$$G_{fNRG}^L = [C_3^I \wedge (U_e^I \times U_i^I)] \tag{109}$$

which is a group of order 12, larger than the *exact* full NRG (108). Rearranging the external and internal motion subgroups we have:

$$G_{fNRG}^L = [C_3^I \wedge U_e^I] \times U_i^I \tag{110-a}$$

The *local* full NRG appears then as a direct product of two commutable subgroups: the restricted NRG corresponding to the internal motions, U_i^I, and the external rotation symmetry subgroup corresponding to the external rotation. The restricted NRG is seen to be smaller than the *exact* full NRG, and the external one isomorphic to the symmetry point group of the molecule in its most symmetric configuration, C_{3v}, in which the hydroxylic bonds are colinear with the C-O bonds.

As a result, the *local* full NRG (110-a) may rewritten as:

$$G_{fNRG}^L \sim U_i^I \times C_{3v} \tag{110-b}$$

expression which exhibits again the form of equation (20).

It may be concluded that, when the external and internal motions may be considered separely, the *local* full NRG is isomorphic to the direct product of the restricted NRG by the symmetry point group of the molecule in its most symmetric configuration (20). This conclusion holds in the case of symmetric linear molecules, in which the reference frame is an axis, except that double valued groups have to be used, but it does not hold in the case of centrosymmetric molecules in which the reference frame is a single atom.

10.APPLICATIONS

Next we shall briefly review some of the numerous applications of the Group Theory for Non-rigid Molecules. These applications are carried out not only for simplifying the numerical calculations, but also for classifiying the solutions, i.e., for interpreting phenomena in order to understand better the physical reality. The comparison between full, complete and local groups gives us a good example how Group Theory for Non-rigid Molecules enables us to distinguish the different types of terms in the Hamiltonian operator.

10.1. Potential Energy Function Determination. The Minimal Expansion

To solve the Schrödinger equation of a non-rigid system, the potential energy function has to be known. This may be expanded in multidimensional Fourier series [30]. Group Theory for Non-Rigid Molecule permits this type of expansions to be simplified and gives rise to approximated analytical forms [30-37] and [47-52].

When internal rotations are studied, these expansions converge quickly so that only a few terms have to be taken into account [35]. When wagging or bending modes are considered larger expansions are usually necessary [51].

In the present work, simple analytical forms were deduced for the potential energy functions of single rotation in phenol (22), of double rotations in benzaldehyde (25) pyrocatechin (30), and acetone (47), of double rotation and inversion in non-planar pyrocatechin (41) and pyramidal acetone(52). Approximate potential energy functions were proposed for some of these non-rigid systems in the local Hamiltonian approach (75), (80), (84), (86), (90) and (96).

10.1.1. The minimal expansion

The use of minimal expansions for an approximate proper description of potential energy functions was proposed few years ago for two-rotor molecules [35]. This kind of expansion retains, in addition to the first harmonics, all the terms necessary for describing the symmetry properties of the potential energy function. This description may be easily extended to many-rotor systems.

In this minimal expansion, two main factors have to be taken into account:

a) The *sine* × *sine* terms describe the cog-wheel effect between the rotors;

b) The first *sine* × *sine* term is maximum when $\theta = 90°$, i.e., when the first harmonic in *cosine* × *cosine* reaches its maximum absolute value. As a result, the first harmonic in *cosine* × *cosine* has to be taken into account for describing properly the shape of the potential energy function, when one *sine* × *sine* term is retained in the expansion for describing the cog-wheel effect.

As a result, the minimal expansion for pyrocathechin will have the simple form, when account is taken of (30):

$$V(\theta_1, \theta_2) = \sum_{K)L=0}^{2} \sum^{2} (\cos K\theta_1 \cos L\theta_2 + \cos L\theta_1 \cos K\theta_2) + A_{11}^{SS} \sin\theta_1 \sin\theta_2 \tag{111}$$

This expansion contains only seven terms. This expression means that only seven conformational energy values have to be determined for retaining all the main features of potential energy surface of pyrocatechin. These seven conformations have to be conveniently chosen in order to avoid linear dependency. In particular, the 90°, 90° and 90°, −90° conformations have to be considered in order to take accurately into account the cog wheel effect. This particularity is often forgotten by the spectroscopists when they determine the V_{11}' term, i.e, the A_{11}^{SS} of the potential energy function (111).

In the same way, the minimal expansions are seen to be built, in the case of rigid pyramidal acetone, or dimethylamine, with only ten terms, using (59) [35,49]:

$$V(\theta_1, \theta_2) = \sum_{K)}^{2} \sum_{L=0}^{2} A_{KL}^{CC} (\cos 3K\theta_1 \cos 3L\theta_2 + \cos 3L\theta_1 \cos 3K\theta_2) +$$

$$+ \sum_{K)0}^{2} A_{K1}^{CC} (\cos 3K\theta_1 \sin 3\theta_2 - \sin 3\theta_1 \cos 3K\theta_2) + A_{11}^{SS} \sin 3\theta_1 \sin 3\theta_2 \tag{112}$$

In reference [35], all the possible two-rotor C_{3v} systems are considered.

These two examples illustrate clearly how the group Theory for Non-rigid molecules can be advantageously used for interpreting and simplifying the potential energy function calculations, without significant decreasing their quality. When the expansion does not converge sufficiently, higher harmonics may be taken into account.

10.1.2. Potential energy calculations

In the potential energy calculations, the large amplitude motions are considered separately from the short amplitude vibration modes. Usually, the

geometry of each conformations is full optimized, i.e., the vibration modes are supposed to follow the large amplitude motions just as the electrons follow the nuclei in the Born-Oppenheimer approximation. This assumption can be very poor.

The different energy values are fitted to a symmetry adapted functional form as described before by a least square procedure when geometry is optimized at each conformation, the rotation angles lose their simple meaning, but the dynamical symmetry of the molecule remains, even when this molecule appears to be deformed during the optimization process.

10.2. Schrödinger Equation Solutions

To solve the Schrödinger equation corresponding to intramolecular conversion motions of a non-rigid molecule, it is convenient to develop the wavefunctions on the basis of the single, double, triple, etc.. free rotor solutions, i.e., on the basis of products of trigonometric functions. The expansion coefficients are then determined by diagonalizing the Hamiltonian matrix constructed with the interaction integrals between the different rotational configurations.

In order to obtain the exact solutions, at least for the lowest energy levels, large expansions have to be considered. The length of these expansions increases with the power of the degree of freedom considered, as a result the calculations could be impossible in some cases, in spite of large computer facilities. Now, if the solutions are developed on the symmetry eigenvectors, the Hamiltonian matrix is factorized, as it is well known, into boxes. The diagonalization of each of these boxes is then much easier from the numerical point of view.

In the present paper, symmetry eigenvectors which factorize the Hamiltonian matrix into boxes are given for the single rotation in phenol (24) for double rotations in benzaldehyde (29), pyrocatechin (34) and acetone (44-46), for double rotation and inversion in non-planar pyrocatechin (40) and pyramidal acetone (49-51). In the same way, symmetry eigenvectors deduced in the local approach are deduced for some of these non-rigid systems (79), (83), and (89). Symmetry eigenvectors for the double internal C_{3v} rotation in molecules with frame of any symmetry are given in reference [36].

10.2.1. Rigid rotor approximation

In order to illustrate the power of the Group Theory for Non-Rigid Molecules, let us consider the double internal rotation problem in acetone, solved in [34]. This motion is described by a restricted Hamiltonian operator such as that of pyrocatechin (32) in which only the threefold periodicity of the potential for acetone (47) has been introduced:

$$H = [-\frac{\partial}{\partial\theta_1}B_1\frac{\partial}{\partial\theta_1} - \frac{\partial}{\partial\theta_1}B_{12}\frac{\partial}{\partial\theta_2} - \frac{\partial}{\partial\theta_2}B_{21}\frac{\partial}{\partial\theta_1} - \frac{\partial}{\partial\theta_2}B_2\frac{\partial}{\partial\theta_2}]+$$
$$+\sum_K\sum_L[A_{KL}^{CC}(\cos 3K\theta_1\cos 3L\theta_2 + \cos 3L\theta_1\cos 3K\theta_2)+ \qquad (113)$$
$$+A_{KL}^{SS}(\sin 3K\theta_1\sin 3L\theta_2 + \sin 3L\theta_1\sin 3K\theta_2)]$$

In this expression B_i are the rotation constants of the rotors and B_{ij} the kinetic interaction term between them. Both terms are written as [53]:

$$B_1 = \frac{\hbar^2}{2}\frac{I_2}{I_1I_2 - \Lambda_{12}^2}, \qquad B_{12} = \frac{\hbar^2}{2}\frac{\Lambda_{12}}{I_1I_2 - \Lambda_{12}^2} \qquad (114)$$

where I_i is the reduced moment of inertia of one of the rotor, and Λ_{ij} the coupling term defined as follows:

$$I_1 = A_1(1 - \sum_{i=a,b,c} A_1\lambda_{1i}^2/M_i), \qquad \Lambda_{12} = \sum_{i=a,b,c} A_1A_2\lambda_{1i}\lambda_{2i}/M_i \qquad (115)$$

In these expressions, A_1 is the moment of inertia of the rotor 1, M_i the three moments of inertia of the whole molecule, and λ_{1i} and λ_{2i} the direction cosines between the internal rotation axes and the ith principal axis.

If we develop the torsional solutions of equation (113) on the basis of double products of trigonometric functions (built up with 16 cosine and 15 sine ones), the Hamiltonian matrix to be diagonalized would be of order $31 \times 31 = 961$.

In contrast, if we resort to the NRG, the Hamiltonian matrix is drastically simplified. As it can be seen in Table 7, the G_{36} group contains 9 irreducible representations, 4 nondegenerate, 4 doubly degenerate, and 1 fourfold degenerate representations. Using the symmetry eigenvectors (44-46), the Hamiltonian matrix of order 961 is factorized in 16 boxes: four of dimension 36, 30, 30 and 25, corresponding to the A_n representations, eight of dimension 55, 45, 55, 45 corresponding to the E_n representations, and four of dimension 110 corresponding to the G representation. As a result, the largest box is of order 110, which is easily diagonalized even in a small computer. This factorization in boxes is illustrated in figure 13.

This simple example illustrates clearly how the NRG's could be advantageously used in problems of two, three, or still higher degrees of freedom, in which the order of the matrix to be diagonalized increases rapidly as the power of the number of degrees of freedom.

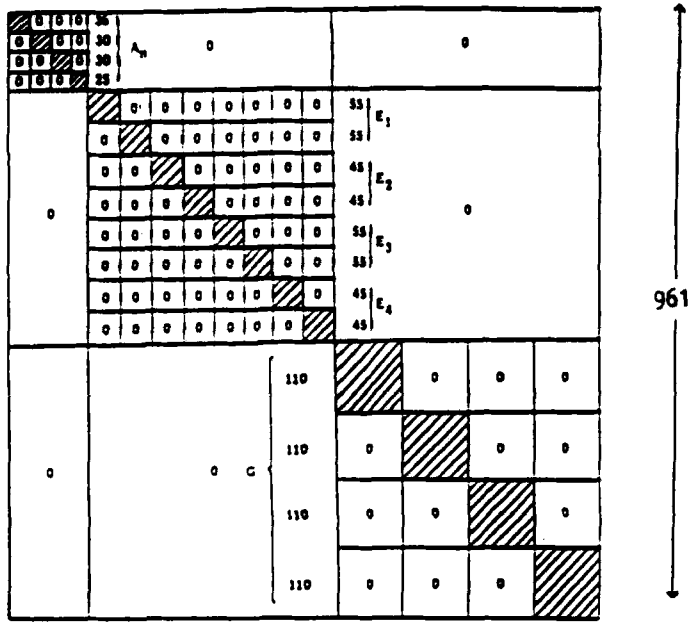


Figure 13. Factorization of the Hamiltonian matrix into boxes for the double rotation in acetone.

We may now to point out that the use of Local Groups, more symmetric than the complete groups, permits a still larger simplification. As an example, the above Hamiltonian matrix for acetone will be factorized in 20 boxes, the largest of which will be of order 60, if the local NRG (95) is used.

10.2.2. Non-rigid rotor approximation

The rigid rotor approximation was used in the previous solution of the Schrödinger equation for the double rotation in acetone [34], i.e., the rotational B constants of the kinetic part of Hamiltonian operator were considered to be constant.

When, the geometry of the molecule is fully optimized at each conformation, the geometry of the whole molecule changes during the rotation. As a result, the rotational constants change with the rotation angles, and have to be fitted in a Fourier series. Since the deformation does not influence the dynamical symmetry properties of the molecule, the rotational B constants may be fitted to a symmetry adapted functional form identical to that of the potential energy function of (113).

10.3 Intramolecular Dynamics

The solution of the Schrödinger equation for the intramolecular motions will furnish us *torsional* levels, and *torsional* wave-functions. The torsional levels will be located more or less deeply in the energy potential energy hypersurface. This hypersurface presents usually potential energy wells, valleys with saddle points, and peaks, which allow some critical points to be defined which enable us to classify the different torsional states.

When the state considered is located above all the energy maxima, no relationship will exist between the phases of the rotation angles of the different rotors, and the rotation may be considered as *free*. In constrast, when the state considered is located below the maxima, but above some saddle points, some pathways will be energetically forbidden on the hypersurface. As a result, some *correlation* appears between the different rotation angles. The lower the torsional levels are located in the potential energy valleys, the more pathways will be forbidden and larger will be the correlation between the rotation angles. Finally, when the torsional state is located into the potential energy well below any saddle point, even the *correlated* rotations will be impossible, and the set of the torsional angles will define a geometry structure, although vibrational motions around the equilibrium positions are always allowed.

10.3.1. Dynamical description of acetone

The torsional levels were determined easily in the two-dimensional case of acetone following the technique described here, as well as in [34]. For this purpose, the energy values of seven conformations adequately chosen (0°, 0°; 0°, 60°; 60°, 60° ; 0°, 30°; 0°, 60°; 30°, 60°; and 30°, −30° were determined by fully optimization into the RHF 6-31 G** approach. These seven values were fitted to the symmetry adapted analytical form (111). The following expression, in cm^{-1}, were obtained:

$$V(\theta_1,\theta_2) = 314.445 - 192.520(\cos 3\theta_1 + \cos 3\theta_2) + 67.555(\cos 3\theta_1 \cos 3\theta_2) +$$
$$+0.197(\cos 6\theta_1 + \cos 6\theta_2) + 1.540(\cos 6\theta_1 \cos 3\theta_2 + \cos 3\theta_1 \cos 6\theta_2) -$$
$$-0.435(\cos 6\theta_1 \cos 6\theta_2) - 83.906(\sin 3\theta_1 \sin 3\theta_2) \qquad (116)$$

The torsional solutions are then determined variationally by using this potential (116), together with B constants (114) evaluated in the rigid rotor approximation, and by developing the solutions on the basis of 961 double products of trigonometric functions.

When the torsional states are weighed with a Boltzmann distribution, it is found that, at 25°C, 46.1% of the acetone molecule are in vibrational *states*,

49.3% in roto-vibrational *states*, in which the rotors are rotating in coupled way and simultaneously vibrating with respect to each other. Finally, 4.6% are in pure rotational *states*. This classification is illustrated in figure 14 as a function of temperature.

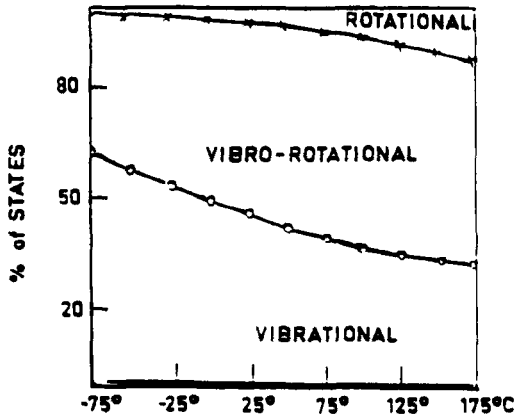


Figure 14. Percentages of vibrational roto-vibrational and pure rotation states in acetone as a function of temperature.

As a result, the acetone is seen to be a relatively floppy system, with a poorly defined structure. This result suggests a graphic representation of the conformational probability density.

10.3.2. Conformational probability density map

The concept of conformational Probability Density Map often is improperly associated with that of potential energy map. The potential energy map indeed gives a relatively good image of the probability density map of finding a molecule in a given conformation [53,54]. To determine the conformational chemical physic properties of a molecule, as a function of temperature, statistical methods have to be used.

For this purpose, two procedures are essentially available [54]:

a) The semi-classical approach;
b) The quantum mechanical approach.

a) Semi-Classical Approach

In the semi-classical approach each point of the potential energy surface is regarded as a *state*, weighed by a Boltzmann distribution. The probability

density is then written as:

$$P(\theta_1, \theta_2) = \frac{exp[-V(\theta_1, \theta_2, ...)/RT]}{\sum_k \sum_l Exp[-V(\theta_k, \theta_l, ...)/RT]} \tag{117}$$

b) **Quantum Mechanical approach**

In the quantum mechanical approach, the density functional of each state weighed by a Boltzmann distribution is taken into account. This density functional is defined as the square of the torsional wave-function. The quantum mechanical probability density is written as:

$$P(\theta_1, \theta_2) = \frac{\sum_n [\sum_i \sum_j C_i^n C_j^n f_i(\theta_1, \theta_2, ..) f_j(\theta_1, \theta_2, ..)] exp(E_n/RT)}{\sum_n \sum_k \sum_l [\sum_i \sum_j C_i^n C_j^n f_i(\theta_k, \theta_l) f_j(\theta_k, \theta_l, ..)] exp(E_n/RT)} \tag{118}$$

where $\sum_i C_i^n f_i(\theta_1, \theta_2, ..)$ are the torsional solutions expressed on the basis set $\{f_i(\theta_1, \theta_2, ..)\}$.

The calculations in the classical approach are indeed much simpler than those performed in the quantum mechanical approach. But they are somewhat inconsistent. First of all, they depend on the number of points considered. In addition, they do not depend on the reduced moment of inertia of the rotor, through the kinetic operator, so that no isotopic effect can be taken into account.

In the semi-classic approach, Group Theory for Non-rigid molecules allows to select the isoenergetic conformations, and to save calculation time.

In the quantum mechanical approach, Group Theory permits us to handle symmetry species, especially by developing the solution on the base of the symmetry eigenvectors. This produre would shorten appreciably the sums on i and j in (118).

In reference [53], conformational probability density maps are given as an example of specific applications of the NRG. These maps are calculated in the quantum mechanical approach for standard and fully deutered acetone molecules and for different temperatures.

The main conclusion put forward in that paper is that deuteration do not influence sensibly the conformation [53]. The probability density maps found for the standard and deutered species look indeed very similar. The second conclusion is that the semi-classical and quantum mechanical approaches yield also similar results [54]. In figures 15 and 16, the conformational probability/$(30°)^2$ density maps for the double rotation in acetone and fully deutered acetone are shown. These maps were calculated at $-75°C$ and $25°C$ using the torsional solutions obtained with the geometry and the potential energy function (116) determined by full optimization.

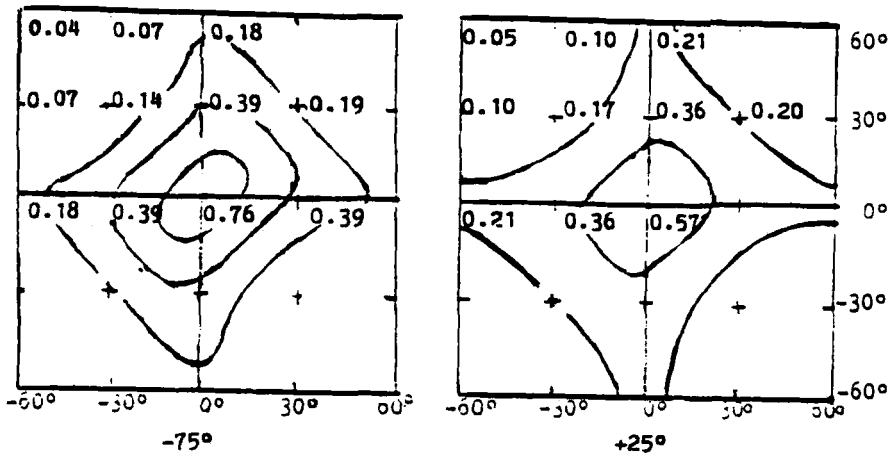


Figure 15. Conformational probability/(30°)² density maps for the double rotation in acetone, at −75°C and 25°C.

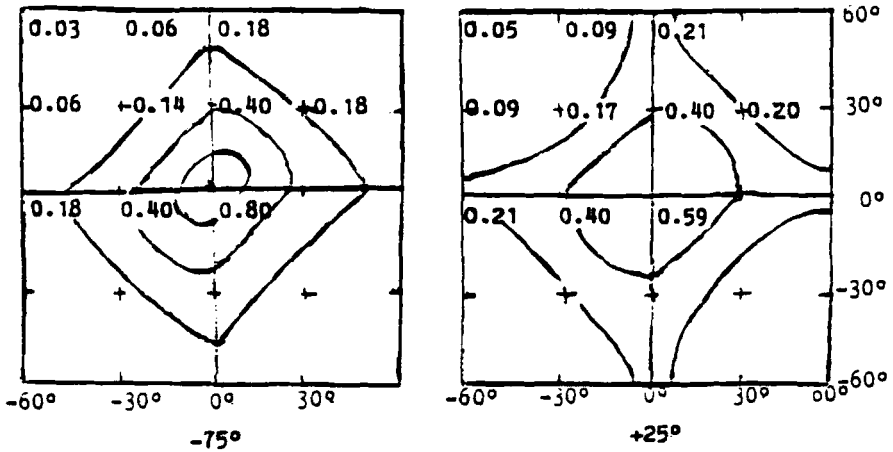


Figure 16. Conformational probability/(30°)² density maps for the double rotation in deutered acetone, at −75°C and 25°C.

As it is seen, such conformational probability density maps depend on temperature and slightly on the isotopic mass of the rotors. As to be expected. the higher the temperature, the smoother are the surfaces. The deuteration does not affect sensibly the conformation. The deutered specie, however, looks slightly more localized. The probability density functions can be used to determine molecular parameter dependencies on temperature [53].

10.4. Theoretical Determination of the Torsional Far Infrared Spectrum of Acetone

Internal rotation in two threefold symmetric top molecules has been mostly studied by microwave spectroscopy [47,55]. Some studies, however, exist in Far Infrared spectroscopy (FIR) [56,57]. Next, the band structure calculations of FIR of acetone will be developed as an example [58,59].

The solution of the Schrödinger equation for the nuclear motions will provide energy levels and torsional wave-functions. Band locations will be given by energy differences and intensities by the square of the electric dipole variation with the transition, weighed by the difference of populations.

$$I \sim (C_i - C_f)\langle \Psi_i | \mu(\theta_1, \theta_2) | \Psi_f \rangle^2 \tag{119}$$

where C_i and C_f are the populations, Ψ_i and Ψ_f the torsional functions, the indexes i and f stand for the initial and final states.

10.4.1. Selection rules

Group Theory for Non-Rigid Molecule (NRG), permits us to classify the torsional wave-functions according to the irreducible representations of the symmetry group of the molecule. As it is well known, the scalar product of (119) does not vanish when the direct products of the irreducible representations, under which Ψ_i and Ψ_f transform, contain at least one of the components of the dipole moments variation. Thus, when symmetry properties of these components are known, **Selection Rules** may be established.

In single rotor molecules such as toluene, phenol or F-phenols [31-32] where the NRG's coincide with some symmetry point groups, the deduction of the selection rules are straightformard. In double rotor molecules, this deduction is not so easy [22]. Let us consider again the acetone case to illustrate this application of the NRG's [58,59].

As well known, the NRG of acetone is the G_{36}. The first step should be to determine the representations of the G_{36} group under which the electric moment components transform. For that purpose, let us consider acetone in a molecular axis system, in which the z axis coincides with the C_2 axis of the molecule, the y axis lies perpendicular to z axis in the CCC plane, and the x axis stands perpendicular to the CCC plane.

With this definition, it is seen that the angle exchange operation $\hat{W}$ is equivalent to a C_2 rotation of the torsional angles, the $\hat{V}$ operation is equivalent to a reflexion in the zy plane, and the $\hat{W}\hat{V}$ operation to a reflexion in the zx plane.

Let us now superimpose the non-degenerate representations of the G_{36} group (or those of the G_4 pyrocatechin group) with their characters, with

the representations of the symmetry point group of molecule, C_{2v} and their characters. This operations is done in Table 14.

Table 14: Superimposition of the pyrocatechin G_4 NRG (capital letters) and the C_{2v} symmetry point (small letters) character tables, with their representations in trigonometric basis sets, as well as the electric dipole moment components

		E	C_2	xz	zy	C_{2v}	G_4	μ
		E	W	WV	V	repre.	representations	
a_1	A_1	1	1	1	1	$cos2\theta$	$\cos\theta_1 + \cos\theta_2$	z
a_2	A_3	1	1	-1	-1	$\sin 2\theta$	$\sin\theta_1 + \sin\theta_2$	R_z
b_1	A_2	1	-1	1	-1	$\cos\theta$	$\sin\theta_1 - \sin\theta_2$	x
b_2	A_4	1	-1	-1	1	$\sin\theta$	$\cos\theta_1 - \cos\theta_2$	y

In the same table the representations of the G_4 and C_{2v} groups, in trigonometric basis sets are given, as well as the electric dipole moment components. So, it is seen that the z, x and y electric dipole components transform according the A_1, A_2 and A_4 representations of the G_4 group, respectively.

A few words has to be said now on the remarkable inversion order observed in the A_n representations, when compared with the C_{2v} ones. This inversion is due to historical reasons when Longuet-Higgins gave for the first time in his historic paper [5] the G_{36} character table for ethane.

From the G_{36} NRG character table, (Table 8), and Table 14, **Selection Rules** for the electric dipole transitions in the Far Infrared Spectrum of acetone may be easily deduced. They are given in Table 15.

10.4.2. Far infrared spectrum of acetone

In Table 15, many transitions appear to be allowed. We have to consider, however, that acetone may be regarded, at least in a first approach, as an oblate symmetrical top, with the largest angular momentum, I_c, parallel to the x axis. In such circunstances, the overall rotation selection rules will be $\Delta K = 0$, for the transition along the x axis, and $\Delta K = \pm 1$, for the transitions in the zy plane. As a result, the transition along the x axis will give rise the a sharp band of C type, while the transitions in the zy plane will give rise to broad flat bands. Up to now, only the sharp bands of C type can be experimentally solved. So, in the following, only the transition along the x axis will be considered, i.e., those of the column on the right of Table 15.

Table 15: Selection Rules for the electric dipole transitions in the FIR spectrum of acetone.

$\mu_z(A_1)$	$\mu_y(A_4)$	$\mu_x(A_2)$
$A_1 \leftrightarrow A_1$	$A_1 \leftrightarrow A_4$	$A_1 \leftrightarrow A_2$
$A_2 \leftrightarrow A_2$	$A_2 \leftrightarrow A_3$	$A_3 \leftrightarrow A_4$
$A_3 \leftrightarrow A_3$		
$A_4 \leftrightarrow A_4$		
$E_1 \leftrightarrow E_1$	$E_1 \leftrightarrow E_1$	$E_1 \leftrightarrow E_1$
$E_2 \leftrightarrow E_2$	$E_3 \leftrightarrow E_4$	$E_2 \leftrightarrow E_2$
$E_3 \leftrightarrow E_3$		$E_3 \leftrightarrow E_4$
$E_4 \leftrightarrow E_4$		
$G \leftrightarrow G$	$G \leftrightarrow G$	$G \leftrightarrow G$

The torsional far infrared spectrum of acetone has been recently studied by using the symmetry point groups [56]. Intensities, however, were not calculated.

The FIR spectrum of acetone, in gas phase, has been studied by several authors [60,61]. Two peaks with some substructure are observed around 125 cm^{-1} and 105 cm^{-1}. Using the potential energy fuction (114), as well as the geometry obtained in a full optimization calculation, two clusters of transitions of appreciable intensities are obtained theoretically. The first one occurs between the torsional microstates $A_1 \rightarrow A_2$, $G \rightarrow G$, $E_1 \rightarrow E_1$ and $E_3 \rightarrow E_4$. The second cluster corresponds to an overtone between the microstates $E_1 \rightarrow E_1$, $E_4 \rightarrow E_3$, $G \rightarrow G$ and $A_2 \rightarrow A_1$.

In the fully deutered acetone, three peaks are essentially observed, and three clusters of transitions of appreciable intensities are theoretically obtained.

Both clusters of transitions are in very good agreement with the experimental values. In addition, the substructure of the observed peaks may be clearly attributed to the existence of torsional microstates, although the external rotational must also contribute.

This result was obtained in the rigid rotor approximation, i.e, with B_1 and B_{12} constant in equation (111). When non-rigidity will be introduced in the calculations, the B constants have to be developed also in Fourier expansions just as the potential. Better results may then be expected [59].

10.5. Torsional Band Structure Determination of Electronic Spectra

The solution of the Schrödinger equation for the nuclear motion furnishes the vibrational levels and wave-functions corresponding to a given electronic state, usually the ground state. If we further consider electronic excited states, band structures of electronic spectra may be determined. As well known, the square of the overlap between the vibrational functions of both states will give us the so-called **Franck-Condon** factors which are a measure of the intensity of the transition between the vibrational levels of both states

$$I_{if} \approx \langle \Psi_i | \Psi_f \rangle^2 \tag{120}$$

The Franck-Condon factors determination is of special interest when the two electronic states, involved in the transition exhibit very different geometries. This is especially the case of electronic transition in the valence shell such as $n \rightarrow \pi^*$, which induces conjugation change, as well as geometrical change, in the molecular system. This phenomenon was studied in the fluorescence spectra of acetaldehyde and acetone [62,63], and in the phosphorescence spectra of thioacrolein and thioacetaldehyde [64,65] and thioacetone [66].

10.5.1. Torsional band structure of the $a^3A'' \leftarrow X^1A'$ of thioacetaldehyde

In the following we shall deduced theoretically the $a^3A'' \leftarrow X^1A'$ torsional band structure of thioacetaldehyde as an example. Thioacetaldehyde exhibits a planar *trans* preferred conformation in the singlet ground state, whereas it presents a non-planar *gauche* ones in the first triplet excited state ($n \rightarrow \pi^*$). The principal deformations involve the torsion of the methyl group and the wagging of the aldehydic hydrogen atom. So, these two variables will be considered in the following.

To classify the *torsional* states of both electronic states with different structures according to the irreducible representations of a same group seems to be difficult. It can be easily done, however, into the NRG formalism. So, the torsional states of both electronic states of thioacetaldehyde can be classified into the representations of a NRG which presents the following characteristic:

a) A periodicity three of the methyl group;

b) Invariance of the energy with respect to a sign change of both rotation and wagging:

$$G_6 = (C_3^I \wedge V^I) \sim C_{3v} \tag{121}$$

which the NRG of random formic acid (63), with the Hamiltonian operator (62).

At this point, we have to solve separately two Schrödinger equations, corresponding to the Hamiltonian operator (62), for both the singlet ground state and the first triplet excited state.

For that purpose the potential energy functions are determined, at 4-31 G levels + d orbitals on the sulphur atom, into the RHF and UHF approximations, for the singlet and triplet states, respectively [51]. In addition, the B constants are calculated separately into the rigid rotor approximation, taking into account the geometrical changes.

The differences between the eigenvalues of both equations will furnish the relative locations of the torsional bands, whereas the Franck-Condon factors (120) will give the relative intensities. Notice, that the nuclear statistical weights have to be taken into account to compare the intensities of transitions belonging to different degrees of degeneracy.

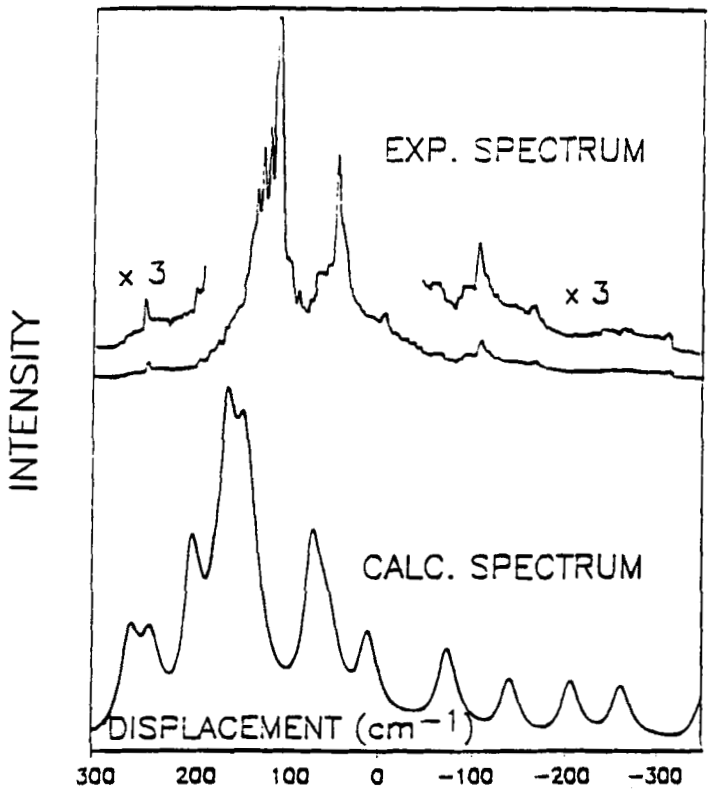


Fig. 17. Band structures corresponding to the methyl torsion and aldehydic hydrogen wagging modes in the $a^3A'' \leftarrow X^1A^1$ spectrum of thioacetaldehyde. The broadness of the bands of the theoretical spectrum is simulated by a Lorentzian profile [67].

The band structures corresponding to the methyl torsion and hydrogen wagging modes in the $a^3A'' \leftarrow X^1A'$ spectrum of thioacetaldehyde is given in figure 17, together with the experimental spectrum. A good agreement is observed [67]. The spectra for the partially and fully deutered species are given in the same reference.

10.5.2. Torsional band structure of the $A^3A_2 \leftarrow X^1A_1$ of Thioacetone

Thioacetone exhibits a planar eclipsed-eclipsed structure in the singlet ground state, whereas it is expected to present a pyramidal gauche-gauche structure in its first triplet excited state. The deformation involves now both methyl rotations and the inversion of the molecule. So, it is a three-dimensional problem [37].

To classify the *torsional* states of both electronic states, the irreducible representations corresponding to the rNRG, G_{36} for pyramidal acetone (48) has to be used:

$$G_{rNRG} = (C_3^I \times C_{3'}^I) \wedge [(WV)^I \times (VU)^I] \tag{122}$$

At this point, the two Schrödinger equations for both the singlet ground state and the first triplet excited state have to be solved in three dimensions. For that purpose, potential energy functions have to be calculated and are in progress in our laboratory.

A preliminary study, however, is already worked out taking into account that the hot band structure depends only on the ground state potential. Taking advantage that the molecule is planar in the ground state, this part of the calculations can be performed in two dimensions. Using a potential calculated as previously in the RHF approximation at the 4-31 G level + d orbitals on the sulphur atom, the first frequency intervals of 148.72/122.24 cm^{-1} for the lowest A_1 symmetry level was found for thioacetone, and fully deutered thioacetone, respectively, in good agreement with the experimental values 153.2/114.7 cm^{-1} [68].

11. CONCLUDING REMARKS

In the present work, we pretend to offer a general description of Group Theory for Non-rigid Molecules. In particular, we have reviewed the Longuet-Higgins' and Altmann's points of view, and we proposed a new theory, more gereral, we call the Non-Rigid Molecule Group (NRG). In contrast to the Longuet-Higgins Theory which resorts to permutations and permutations-inversions, this new theory retains the Schönfliess formalism of physical op-

erations, just as Altmann does, in order to conserve an easy physical interpretation.

The concepts of full and restricted NRG's are defined. The full NRG's, which consider the overall rotations, are seen to be entirely equivalent to the Molecular Symmetry Groups of Longuet-Higgins [5], whereas the restricted NRG's, limited to the interconversion motions, may be compared with the Isodynamic groups of Altmann, and the isometric groups of Günthard [10,27].

In order to introduce the NRG formalism, the necessary mathematical machinery is introduced in a very simple and gradual way considering a lot of examples. Finally, the concept of local group, which can not be introduced easily into the Longuet-Higgins formalism, is proposed for simplifying still more the very involved non-rigid problems and classifying in more detail their solutions.

Throughout this paper, we give potential energy functions, symmetry eigenvectors, as examples, for systems of one, two and three internal degrees of freedom, in the formalism of the restricted Hamiltonian operator, as well as in the local one. A generalization of these ideas can be found in the scientific literature [21,22] and [30-37].

Finally, the main applications of the rNRG's are reviewed. So, for example, symmetry adapted minimal expansions are proposed for determining potential energy functions. Symmetry eigenvectors which factorize the Hamiltonian matrix in boxes are given to construct the torsional solutions of the Schrödinger equation for the nuclear motion. Internal dynamics of non-rigid systems is studied in the case of acetone, and conformational probability density map are drawn by using symmetry eigenvectors [53,54]. Selection rules for torsional transitions in the far-infrared region, as well as the most relevant features (C bands) of the FIR acetone spectrum were determined theoretically [58,59]. Finally, torsional band structures in the triplet-singlet spectra of thioacetaldehyde and thioacetone are determined, classiflying the *torsional* states of both electronic states into a same NRG [68].

Through all this study, references are given to the most relevant papers of the field. We do not pretend to gather here all the bibliography that should be too voluminous and can be found in the citated papers.

Ackowledgements

The author is endebted to Prof. David C. Moule for improving the English editing of this paper, to Prof. Jean Brocas for reading the manuscript, and to Prof. James K.G. Watson for some valuable comments.

Bibliography

[1] J.D.Swalen and C.C.Costain, J.Chem.Phys.,**31**, 1561 (1959).

[2] R.J.Myers and E.B.Wilson, J.Chem.Phys.,**33**, 186 (1960).

[3] von H. Dreizler, Z.Naturforsch.,**16a**, 477 (1961); ibid.,**16a**. 1354 (1961).

[4] J.T.Hougen, J.Chem.Phys.,**37**, 1433 (1962); ibid.**39**, 358 (1963).

[5] H.C.Longuet-Higgins, Mol.Phys.,**6**, 445 (1963).

[6] A.J.Stone, J.Chem.Phys.,**41**, 1568 (1964).

[7] R.S.Dalton, Mol.Phys.,**11**, 265 (1966).

[8] P.R.Bunker, in *Vibrational Spectra and Structure*, Durig ed., Marcel Dekker, New York,1975.

[9] P.R.Bunker, in " *Molecular Symmetry and Spectroscopy*". Academic P., New York, 1979.

[10] S.L. Altmann, Proc.Roy.Soc., **A298**, 184 (1967).

[11] S.L. Altmann, in *Induced Representations in Crystals and Molecules.*, Academic Press, London, 1977.

[12] R.B. Woodward and R. Hoffmann, J. Am. Chem. Soc.. **87**, 395, 2045, 2511 (1965); in *The Conservation of Orbital Symmetry*.. Verlag Chemie, Berlin, 1970.

[13] R.E. Stanton and J.W. McIver,Jr., J.Am.Chem.Soc., **97**. 3632 (1975).

[14] Y. Ellinger and J. Serre, Int. J.Quantum Chem.,**S7**, 217 (1973).

[15] T.D. Bauman, C.D. Duncan and C. Trindle, Int.J. Quantum Chem., **11**, 399 (1977).

[16] M. Quack, Mol. Phys., **34**, 477 (1977).

[17] A. Janner and T. Janssen, Phys.Rev.,**B-15**, 643 (1977).

[18] A. Toro-Labbé and J. Maruani, in *Structure and Dynamics of Molecular Systems-II.*, Daudel et al. eds., Reidel, Dordrecht, 1956. pp.35-56; J. Maruani and A. Toro-Labbé, Can. J.Chem., **66**, 1948 (1988).

[19] Y.G. Smeyers, Fol.Chim. Theoret.Lat.,**6**, 139 (1978).

[20] C.M. Woodman, Mol. Phys.,**19**, 753 (1970).

[21] Y.G. Smeyers," *Teoría de Grupos para Moléculas No-rígidas. El Grupo Local. Aplicaciones*", in Memorias, Vol.13, Real Academia de Ciencias, Madrid, 1989.

[22] Y.G. Smeyers, Fol.Chim. Theoret. Lat., **17**, 15 (1989).

[23] M.Born and J.R.Oppenheimer, Ann. Physik, **84**, 457 (1927).

[24] P.H.E. Meier and E. Bauer," *Group Theory*", North-Holland, Amsterdam,1962.

[25] J.M.F. Gilles and J. Philippot, Int. J.Quantum Chem., **6**, 225 (1972).

[26] G.S. Ezra, Mol. Phys.,**38**, 863 (1979).

[27] H. Frei, A Bauder and Hs.H. Günthard," *Large Amplitude Motion in Molecules*",in Topics in Current Chemistry, Vol. 81, Springer (1979).

[28] G.A. Natanson," *On the Invariance of the Localized Hamiltonians under Feasible Elements of the Nuclear Permutation-Inversion Group*" in Advances in Chemical Physics, Vol.58, Prigogine Ed., John Wiley,New York, 1985.

[29] S.L. Altmann, Rev. Mod. Phys.,**35**, 641 (1963).

[30] J. Maruani, A. Hernandez-Laguna and Y.G. Smeyers, J.Chem. Phys.,**63**, 4515 (1975).

[31] Y.G. Smeyers, A. Hernandez-Laguna and P. Galera-Gomez, An. Quim., **76**, 67 (1980).

[32] Y.G. Smeyers and A. Hernandez-Laguna, Int.J. Quantum Chem., **22**, 681 (1982); J.Mol. Struct (Theochem),**149**, 127 (1987).

[33] J. Maruani, Y.G. Smeyers, and A. Hernandez-Laguna, J.Chem. Phys.,**76**, 3123 (1982).

[34] Y.G. Smeyers and M.N. Bellido, Int.J.Quantum Chem.,**19**, 553 (1981).

[35] Y.G. Smeyers, J.Mol.Struct. (Theochem), **107**, 3 (1984).

[36] Y.G. Smeyers and A. Niño, J. Comp. Chem.,**8**, 380 (1987).

[37] Y.G. Smeyers, M.N. Bellido, and A. Niño, J.Mol.Struct. (Theochem), **166**, 1 (1988).

[38] J.K.G. Watson, Mol.Phys.,**21**, 577 (1971).

[39] S.L. Altmann, Mol.Phys.,**21**, 587 (1971).

[40] B.J. Dalton, J.Chem.Phys., **54**, 4745 (1971).

[41] J. Brocas and D. Fastenakel, Mol. Phys.,**30**, 193 (1975).

[42] B.J. Dalton, J. Brocas and D. Fastenakel, Mol. Phys., **31**, 1887 (1976).

[43] C. Trindle, S.N. Datta and T.D. Bauman, Int.J. Quantum Chem., **11**, 627 (1977).

[44] J. Brocas, D. Fastenakel and J. Buschen, Mol. Phys., **41**, 1163 (1980).

[45] J. Brocas and C. Rusu, Int. J. Quantum Chem., **22**, 331 (1982).

[46] R.S. Berry, J.Chem.Phys., **32**, 933 (1962).

[47] P.Groner and J.R. Durig, J.Chem.Phys.,**66**, 1856 (1977).

[48] A. Toro-Labbé and J. Maruani, Int.J. Quantum Chem., **22**, 115 (1982).

[49] Y.G. Smeyers and M.N. Bellido, Int.J. Quantum Chem., **23**, 507 (1983).

[50] Y.G. Smeyers, A. Huertas-Cabrera and S. Suhai, Theoret. Chim. Acta, **64**, 97 (1983).

[51] Y.G. Smeyers, A. Niño and M.N. Bellido, Theoret. Chim. Acta, **69**, 259 (1988).

[52] Y.G. Smeyers, M. Fernandez-Nuñez and V. Botella, J. Mol.Struct. (Theochem), **210**, 273 (1990).

[53] Y.G. Smeyers and A. Hernandez-Laguna, in *Structure and Dynamics of Molecular Systems*, Daudel et al. eds., Reidel, Dordrecht, 1985, pp. 23-40.

[54] Y.G. Smeyers and A. Hernandez-Laguna, Int.J. Quantum Chem., **29**, 553 (1986).

[55] J. Meier, A. Bauder and Hs.H. Günthard, J.Chem. Phys., **57**, 1219 (1972).

[56] J.R. Durig, Y.S. Li and P. Groner, J.Mol. Spectrosc., **62**, 159 (1976).

[57] A.G. Ozkabak, J.G. Phillis and L. Goodman, J.Am. Chem.Soc., **112**, 7854 (1990).

[58] Y.G. Smeyers, A. Hernandez-Laguna, M.N. Bellido and A. Niño, Fol.Chim.Theoret.Lat., **16**, 185 (1988).

[59] Y.G. Smeyers, M.L. Senent, V. Botella and D.C. Moule, results to be published.

[60] W.G. Fateley and F.A. Miller, Spectrochim. Acta, **18**, 997 (1962).

[61] P. Groner, G.A. Guirgis and J.R. Durig, J.Chem. Phys., **86**, 565 (1987).

[62] M. Noble and E.K.C. Lee, J. Chem. Phys., **81**, 1632 (1984).

[63] M. Baba, I. Hanazaki and V. Nagashima, J.Chem. Phys., **82**, 3938 (1985).

[64] D.C. Moule, R.H. Judge, H.L. Gordon and J.D. Goddard, Chem. Phys.,**105**, 97 (1986).

[65] R.H. Judge, D.C. Moule, A.E. Bruno and R.P. Steer,J.Chem. Phys., **87**, 60 (1987).

[66] D.C. Moule, Y.G. Smeyers, M.L. Senent, D.J. Clouthier, J. Karolczak and R.H. Judge, J. Chem. Phys., **95**, 3137 (1991).

[67] Y.G. Smeyers, A. Niño and D.C. Moule, J. Chem. Phys. **93**, 5786 (1990).

[68] Y.G. Smeyers, M.L. Senent and D.C. Moule, Int. J. Quantum Chem., **S-24**, 835 (1990).

Some Studies of the General Hartree-Fock Method

by Per-Olov Löwdin* and István Mayer**

Quantum Theory Project
362 Williamson Hall, University of Florida
Gainesville, Florida 32611, USA

* Professor Emeritus at Uppsala University, Uppsala, Sweden.
** Permanent address: Central Research Institute for Chemistry, Hungarian Academy of Sciences, H-1525 Budapest, Hungary.

List of contents:

1. History of the General Hartree-Fock Method

In the independent-particle-model (IPM) originally due to Bohr [1], each particle moves under the influence of the outer potential and the average potential of all the other particles in the system. In modern quantum theory, this model was first implemented by Hartree [2], who solved the corresponding one-electron Schrödinger equation by means of an iterative numerical procedure, which was continued until there was no change in the significant figures associated with the electric fields involved so that these could be considered as self-consistent; this approach was hence labelled the Self-Consistent-Field (SCF) method. In order to take the Pauli exclusion principle into account, Waller and Hartree [3] approximated the total wave function for a N-electron system as a product of two determinants associated with the electrons of

different spins. Somewhat later, Slater [4] introduced the spin coordinate ζ and assumed that the one-electron functions involved $\psi = \psi(x)$ depended on a combined space-spin coordinate $x = (\mathbf{r},\zeta)$ and introduced further the convention that in calculating the binary product $\langle\psi_1 | \psi_2\rangle$ one should integrate over the space and sum over the spin. In treating a N-electron system in the independent-particle model, Slater approximated the total wave function Ψ by a *single determinant* $\Psi \approx D$ formed by the one-electron functions $\psi_1, \psi_2, \psi_3, \ldots\psi_N$, which is antisymmetric and automatically satisfies the Pauli exclusion principle. In the Born-Oppenheimer approximation [5], the total Hamiltonian H for an atomic, molecular, or solid-state system containing N electrons has the form:

$$H = e^2 \sum_{g<h} Z_g Z_h / r_{gh} + \sum_i (\mathbf{p}_i^2/2m - e^2 \sum_g Z_g/r_{ig}) + \sum_{i<j} e^2/r_{ij}$$

$$= H_{(0)} + \Sigma_i H_i + \Sigma_{i<j} H_{ij} , \qquad (1.1)$$

Applying the variation principle $\delta\langle H\rangle = 0$ to the total Hamiltonian H for the N-electron system, Slater [6] and Fock [7] derived effective one-electron equations of the form

$$F(1)\, \psi_k(x_1) = \Sigma_l \psi_l(x_1)\, \lambda_{lk}, \qquad (1.2)$$

which became known as the Hartree-Fock (HF) equations. Slater had pointed out that - except for an irrelevant factor - the determinant D is invariant under a non-singular linear transformation of the one-electron functions $\{\psi_1, \psi_2, \psi_3, \ldots\psi_N\}$, which hence span a subspace of the total one-electron Hilbert space. The Lagrangian multipliers λ_{lk} form an hermitean matrix $\lambda = \{\lambda_{lk}\}$, and it could then be shown that the equations (1.2) by a unitary transformation could be brought to the simple form

$$F(1)\, \psi_k(x_1) = \varepsilon_k \psi_k(x_1) , \qquad (1.3)$$

where the eigenvalues ε_k were interpreted as one-electron energies. There are different types of solutions for various types of Lagrangian multipliers $\lambda = \{\lambda_{lk}\}$ - some are for instance localized - and it should be observed that the so-called *canonical Hartree-Fock functions* defined by (1.3) are always linear combinations of the other solutions. The effective Hamiltonian F(1) is often referred as the Fock-operator. Fock could show that, if the one-electron functions $\psi = \{\psi_1, \psi_2, \psi_3, \ldots\psi_N\}$ are

chosen to be orthonormal, so that $\langle\psi_k|\psi_l\rangle = \delta_{kl}$, then the properties of the system is described by the density matrix

$$\rho(x_1,x_2) = \Sigma_k\, \psi_k(x_1)\psi_k^*(x_2) \qquad (1.4)$$

where $\rho(x_1,x_2)$ is the kernel of an operator ρ having the properties

$$\rho^2 = \rho, \qquad \rho^\dagger = \rho, \qquad \mathrm{Tr}\,\rho = N. \qquad (1.5)$$

The properties of the Hartree-Fock scheme were further investigated by Dirac [8], who showed that the Fock operator may be represented in the simple form

$$F(1) = H_1 + e^2 \int dx_2 \frac{\rho(2,2) - \rho(1,2)\, P_{12}}{r_{12}}, \qquad (1.6)$$

where P_{12} is a permutation or exchange operator defined through the relation $P_{12}\, u(1) = u(2)$. The term containing P_{12} is often referred to as the *exchange term* in the Fock operator. It is then easily shown that the operator F has the property $\langle v|Fu\rangle = \langle Fv|u\rangle$, i.e. that the Fock operator is a self-adjoint one-electron operator. The kernel $\rho(1,2)$ is often referred to as the Fock-Dirac density matrix, and it is evident that the operator ρ is the projector on the subspace spanned by the N one-electron functions. The one-electron energies ε_k have finally been given a simple physical interpretation by Koopmans [9].

The one-electron functions $\psi_k = \psi_k(x) = \psi_k(\mathbf{r},\zeta)$ are not restricted in any way - they are general *spin-orbitals* - and one may hence characterize this approach as the General Hartree-Fock (GHF) method. It should be observed, however, that the development of the Hartree-Fock scheme went in a different direction away from this general approach.. The calculations by Hartree in the years 1928-1933 [10] were carried out without exchange, and the first calculations *including exchange* were carried out on the ground state of the lithium atom by Fock and Petrashen [11]. The Li-atom in its ground state has the symmetry ^2S and is characterized by the configuration $(1s)^2(2s)$, and it seemed hence natural to start the calculation from the configuration $(1s\alpha, 1s\beta, 2s\alpha)$, where α and β are the elementary spin functions. These authors found - certainly somewhat to their surprise - that the calculations contained a *non-diagonal* Lagrangian multiplier $\lambda_{1s,2s}$, which could not be transformed

away as in the general theory. Constants of motion related to *symmetry properties* of the total Hamiltonian are, of course, of fundamental importance in quantum theory, and the question was now how they should be incorporated in the Hartree-Fock scheme. Already in 1930, it had been shown by Delbrück [12] that, if the total system has spherical symmetry and one assumes that the starting determinant has ^1S symmetry, then this assumption is self-consistent, and the final Hartree-Fock functions turn out to be eigenfunctions of the orbital angular momentum and the spin. One could hence expect that the symmetry properties would be a natural part of any Hartree-Fock calculation.

In 1951, Slater [13] returned to this problem and pointed out that, if the Li-atom has the configuration (1sα,1sβ,2sα), then the 1sα–function is influenced by a different exchange potential than the 1sβ-function, since exchange occurs only between electrons with parallell spins. and one could then expect that the calculations would lead to two different 1s-orbitals. Slater called this phenomenon *spin-polarization* or spin-splitting, and expected that it would be a small but important effect. Later applications at M.I.T. to open-shell systems [14] indicated that the effect was anything but small, and that it had many interesting consequences, particularly as to magnetic properties. A more detailed analysis of the new approach showed [15] that the spin-polarization was to some extent coupled with the *correlation splitting* due to the Coulomb repulsion, in which electrons with anti-parallell spins try to avoid each other by permitting "different orbitals for different spins" (DODS).

In the later part of the 1950's, it was evident that it was necessary to distingush the new approach dealing with different orbitals for α-spin and β-spin from the previous approach starting out from symmetry restrictions: the latter was called the Restricted Hartree-Fock (RHF) scheme, whereas the new approach was called the Unrestricted Hartree-Fock (UHF) scheme. For some time there was a certain amount of competition between the two schemes. In the late 1950's, it was further shown that the RHF-scheme for closed-shell systems was completely *self-consistent* not only for atoms but also for molecules and solids [16,17] and that, if one started by imposing a symmetry requirement on the original Slater determinant, this assumption would be self-consistent, i.e. the final determinant would have the same symmetry property. Since symmetry properties are of such fundamental importance in quantum theory, one would hence anticipate that the RHF-scheme would

be the fundamental one. The situation was complicated by the fact that, in many applications, the UHF-scheme would have a lower energy <H> than the RHF-scheme.

In 1963, Löwdin [18] pointed out the occurrence of a *symmetry dilemma* in the Hartree-Fock method: that, even if a symmetry requirement is self-consistent, it is still a *constraint* which will increase the energy <H>, and the associated optimum value of <H> is hence only a *local minimum*; on the other hand, if one looks for the *absolute minimum* of <H>, the associated Slater determinant may very well be a *mixture* of various symmetry types. It is evident that some of the optimum values of <H> are not even local mimima, and the study of the nature of the optimum values by means of the second-derivaties or the *Hessians* has become one of the most intensely studied problems [19] in the current literature. Most of the papers use very elegant and forceful methods based on the use of second quantization, but it should be observed that the problem may also be treated in an elementary way [20].

It should be observed that, even if the UHF-scheme is less restricted than the RHF-scheme, it still deals with pure α- and β-functions, and it is hence restricted in comparison to the General Hartree-Fock (GHF) scheme, which deals with general one-electron functions of the form:

$$\psi_k(x) = \psi_k(\mathbf{r},\zeta) = \psi_{k+}(\mathbf{r})\,\alpha(\zeta) + \psi_{k-}(\mathbf{r})\,\beta(\zeta). \quad (1.7)$$

We note that the GHF-scheme is identical with the Hartree-Fock scheme originally derived by Fock and Slater, since in their derivations there were no restrictions imposed on the nature of the one-electron functions. This means that the GHF-scheme is still based on the equations (1.3)-(1.6). It should be observed that the functions $\psi_{k+}(\mathbf{r})$ and $\psi_{k-}(\mathbf{r})$ in general are of *complex* nature.

It is obvious that the Hartree-Fock equations are rather complicated integro-differential equations of a non-linear nature with bifurcations etc., and it was hence of fundamental importance when Lieb and Simon [21] in 1977 could show the *mathematical existence* of solutions to these equations. There are still some mathematical problems associated with the Hartree-Fock scheme, particularly the connection between the starting point of the calculations and the final result, which is usually associated with a "local minimum" of the energy <H>. We note further that the concept of "self-consistency" is related to some form of "numerical convergence" in a specified number

of significant figures in the calculations and not to the concept of *mathematical convergence*, which is a problem that has so far not been sufficiently investigated. It should also be observed that the SCF-procedure is an iterative procedure which is subject to the laws discovered by Schröder [22] in 1870, which may be used to speed up the convergency or to change an obviously divergent process into a convergent one.

During the last decades, there has slowly been a return of the interest of the scientists working on many-fermion systems from the RHF- and UHF-schemes to the original Hartree-Fock scheme dealing with general spin-orbitals. In 1960, Overhauser [23] showed in a study of a one-dimensional Fermi gas that there existed self-consistent solutions in the form of giant "spin density waves" which had a lower energy than the plane-wave state, and in solid-state theory one has later shown the existence of both "spin-density waves" and "charge density waves", dealing with general spin-orbitals [24]. After the formulation of the "symmetry dilemma", some examples involving negative atomic ions were also given [25]. In nuclear physics as well as in the theory of electronic systems in general, the same problem was studied by Fukutome [26] who also developed a *general classification scheme* for the various general spin-orbitals occuring as solutions to the GHF-equations based on their symmetry properties, and introduced the terms time-reversal-invariant closed shell solutions (TICS), charge-current waves (CCW), axial spin waves (ASW), axial spin-current waves (ASCW), axial spin density waves (ASDW), torsional spin waves (TSW), torsional spin-current waves (TSCW), torsional spin density waves (TSDW), etc.

2. The General Hartree-Fock Equations; Separation of Space and Spin; the MO-LCAO-approach

Some notations. - Before starting the main part of this section, it may be convenient to discuss the use of some specific notations. Dirac [27] considered the bracket $\langle x|y\rangle$ as the scalar product of a bra-vector $\langle x|$ and a ket-vector $|y\rangle$, and - in this connection - he also introduced the ket-bra operators $T = |b\rangle\langle a|$ defined through the relation $Tx = b\langle a|x\rangle$. They satisfy the relations $T^2 = \langle a|b\rangle T$, $T^\dagger = |a\rangle\langle b|$, and $\mathrm{tr}\, T = \langle a|b\rangle$.

In the following we will further use bold-face symbols **A** and **B** to denote rectangular matrices - including row and column vectors and quadratic matrices. If **A** and **B** are rectangular

matrices of order m x p and p x n, we will let the product $\mathbf{C} = \mathbf{A}\,\mathbf{B}$ define a matrix of order m x n, defined through the relation

$$(\mathbf{A}\,\mathbf{B})_{kl} = \Sigma_\alpha\, A_{k\alpha} B_{\alpha l}, \qquad (2.1)$$

i.e. one multiplies the rows of the first matrix in order with the columns of the second matrix. As an illustration, we will consider the row vector formed by the one-electron functions $\psi = \{\psi_1, \psi_2, \psi_3, \ldots \psi_N\}$. The orthonormality property may now be expressed in the condensed form $\langle\psi\,|\psi\rangle = \mathbf{1}$, where as the Fock-Dirac operator may be written in the form $\rho = |\psi\rangle\langle\psi|$, i.e.

$$\rho = |\psi\rangle\langle\psi|, \qquad \langle\psi\,|\psi\rangle = \mathbf{1}, \qquad (2.2)$$

and the validity of the relations (1.5) becomes now clear and transparent in a somewhat new way.

If one instead starts from N one-electron functions $\psi = \{\psi_1, \psi_2, \psi_3, \ldots \psi_N\}$, which are linearly independent but not necessarily orthonormal, one may by means of symmetric orthonormalization [28] introduce an orthonormal set $\varphi = \psi\,\langle\psi|\psi\rangle^{-1/2}$, which gives the projector

$$\rho = |\varphi\rangle\langle\varphi| = |\psi\rangle\langle\psi|\psi\rangle^{-1}\langle\psi|, \qquad (2.3)$$

which has been studied in greater detail elsewhere [29].

The General Hartree-Fock equations. - The one-electron functions in the GHF-scheme are of the form (1.7) or

$$\psi_k(x) = \psi_k(\mathbf{r},\zeta) = u_k(\mathbf{r})\,\alpha(\zeta) + v_k(\mathbf{r})\,\beta(\zeta), \qquad (2.4)$$

where the orbital functions $u_k(\mathbf{r})$ and $v_k(\mathbf{r})$ are in general *complex* functions. Introducing the row vectors $\mathbf{u} = \{u_1, u_2, u_3, \ldots u_N\}$ and $\mathbf{v} = \{v_1, v_2, v_3, \ldots v_N\}$, one may write the relations (1.7) for k = 1,2,3,...N in the condensed form

$$\psi = \mathbf{u}\,\alpha + \mathbf{v}\,\beta. \qquad (2.5)$$

Since the spin functions α and β are orthonormal, one gets directly

$$\langle\psi|\psi\rangle = \langle \mathbf{u}\,\alpha + \mathbf{v}\,\beta | \mathbf{u}\,\alpha + \mathbf{v}\,\beta\rangle = (\mathbf{u}|\mathbf{u}) + (\mathbf{v}|\mathbf{v}), \quad (2.6)$$

where the round brackets indicate binary products in the orbital space. In the following, we will use the notation

$$\mathbf{d} = \langle\psi|\psi\rangle^{-1}, \quad (2.7)$$

where $\mathbf{d}$ is a matrix of order N x N with the elements d_{kl}. According to (2.3), the Fock-Dirac operator takes now the form

$$\rho = |\psi\rangle\langle\psi|\psi\rangle^{-1}\langle\psi| = |\psi\rangle\,\mathbf{d}\,\langle\psi| =$$

$$= \sum_{k\,l=1}^{N} |\psi_k\rangle\, d_{kl}\, \langle\psi_l| \ , \quad (2.8)$$

where the double sum over k and l contains N^2 terms, whereas for the associated density matrix one gets the formula

$$\rho\,(1,2) = |\psi(1)\rangle\,\mathbf{d}\,\langle\psi(2)| = \Sigma_{kl}\,\psi_k(1)\,d_{kl}\,\psi_l^*(2) \ . \quad (2.9)$$

Substituting the expression (2.5) into (2.8), one gets further

$$\rho = |\mathbf{u}\,\alpha + \mathbf{v}\,\beta\rangle\,\mathbf{d}\,\langle\mathbf{u}\,\alpha + \mathbf{v}\,\beta| =$$

$$= \rho^{++}\,\alpha\,\alpha + \rho^{+-}\,\alpha\,\beta + \rho^{-+}\,\beta\,\alpha + \rho^{--}\,\beta\,\beta \ , \quad (2.10)$$

where

$$\rho^{++} = |\mathbf{u})\,\mathbf{d}\,(\mathbf{u}|, \qquad \rho^{+-} = |\mathbf{u})\,\mathbf{d}\,(\mathbf{v}|,$$

$$\rho^{-+} = |\mathbf{v})\,\mathbf{d}\,(\mathbf{u}|, \qquad \rho^{--} = |\mathbf{v})\,\mathbf{d}\,(\mathbf{v}|. \quad (2.11)$$

In this way, it is hence possible to separate the orbital parts and the spin parts of the Fock-Dirac operator in a simple way. The UHF-scheme, which deals with only pure α- and β-functions, is characterized by the fact that one has $\rho^{+-} = \rho^{-+} = 0$, whereas in the GHF-scheme these components are always non-vanishing. According to (1.3), one may now write the Hartree-Fock equations in the form

$$F(1)\,\psi(1) = \psi(1)\,\varepsilon \ , \quad (2.12)$$

where ε is the diagonal matrix having the elements $\varepsilon_k\delta_{kl}$. According to (1.6), the Fock operator has here the form

$$F(1) = H_1 + e^2 \int dx_2 \frac{\rho(2,2) - \rho(1,2)\, P_{12}}{r_{12}} =$$

$$= H_1 + J_1 - K_1 , \qquad (2.13)$$

where J_1 and K_1 are the coulomb and exchange operators in the original Hartree-Fock scheme. This gives us directly

$$J_1 = e^2 \int dx_2 \frac{\rho(2,2)}{r_{12}} = e^2 \int dr_2 \frac{\rho^{++}(2,2) + \rho^{--}(2,2)}{r_{12}} , \qquad (2.14)$$

where the quantity $[\rho^{++}(2,2) + \rho^{--}(2,2)]$ in the numerator of the last integrand is the orbital density $N(2,2)$ of electron 2. For the exchange operator, one gets in a similar way

$$K_1 = e^2 \int dx_2 \frac{\rho(1,2)\, P_{12}}{r_{12}} , \qquad (2.15)$$

and further

$$K_1 \psi(1) = e^2 \int dx_2 \frac{\rho(1,2)\, P_{12}}{r_{12}} \psi(1) = e^2 \int dx_2 \frac{\rho(1,2)\, \psi(2)}{r_{12}} =$$

$$= e^2 \int \frac{dx_2}{r_{12}} [\rho^{++}(1,2)\, \alpha_1\alpha_2 + \rho^{+-}(1,2)\alpha_1\beta_2 +$$

$$\rho^{-+}(1,2)\beta_1\alpha_2 + \rho^{--}(1,2)\beta_1\beta_2] \times [\mathbf{u}(2)\, \alpha_2 + \mathbf{v}(2)\, \beta_2] = \qquad (2.16)$$

$$= e^2 \int \frac{dr_2}{r_{12}} [\rho^{++}(1,2)\mathbf{u}(2)\alpha_1 + \rho^{+-}(1,2)\mathbf{v}(2)\alpha_1 +$$

$$+ \rho^{-+}(1,2)\mathbf{u}(2)\beta_1 + \rho^{--}(1,2)\mathbf{v}(2)\beta_1],$$

where in the last line we have summed over the spin coordinate ζ_2. At this point, it may be convenient to introduce the notation

$$K_1{}^{\mu\nu} = e^2 \int dr_2 \frac{\rho^{\mu\nu}(1,2)\, P_{12}{}^{r}}{r_{12}} , \qquad (2.17)$$

where $\mu,\nu = +,-$, and the permutation operator $P_{12}{}^{r}$ works only on the orbital coordinates. From the last line in equation (2.16), one hence obtains

$$K_1 \psi(1) = K_1{}^{++}\mathbf{u}(1)\alpha_1 + K_1{}^{+-}\mathbf{v}(1)\alpha_1 +$$

$$K_1{}^{-+}\mathbf{u}(1)\beta_1 + K_1{}^{--}\mathbf{v}(1)\beta_1 . \qquad (2.18)$$

In studying the Hartree-Fock equations (2.12), one can now easily separate the α- and β-components, and using (2.13) and (2.18), one obtains directly

$$(H_1 + J_1 - K_1^{++})\,\mathbf{u}(1) - K_1^{+-}\mathbf{v}(1) = \mathbf{u}(1)\,\varepsilon\,, \qquad (2.19)$$

$$-K_1^{-+}\mathbf{u}(1) + (H_1 + J_1 - K_1^{--})\,\mathbf{v}(1) = \mathbf{v}(1)\,\varepsilon\,, \qquad (2.20)$$

and it is then possible to combine these two relations into a matrix equation of the form

$$\begin{Bmatrix} H_1 + J_1 - K_1^{++}; & -K_1^{+-} \\ -K_1^{-+}; & H_1 + J_1 - K_1^{--} \end{Bmatrix} \begin{Bmatrix} \mathbf{u}(1) \\ \mathbf{v}(1) \end{Bmatrix} = \begin{Bmatrix} \mathbf{u}(1) \\ \mathbf{v}(1) \end{Bmatrix} \varepsilon\,. \qquad (2.21)$$

We note that these equations for the 2N orbital functions **u** and **v** are identical to the original Hartree-Fock equations and that no additional assumptions have been made.

The MO-LCAO approach for solving the HF-equations. -

During the first 25 years of the development of the Hartree-Fock method, one used various types of numerical integration in connection with the SCF-procedures for studying atomic structures, and special methods were developed by Hartree [30]. There was also a renewed interest in the forceful central difference methods, which go back to Gauss [31]. However, in order to treat the electronic structure of *molecules* by the Hartree-Fock method, there were evidently some new ideas needed. Already in 1928, it had been shown by Mulliken and Hund [32] that one can obtain a good understanding of molecular structure by means of the concept of *molecular orbitals*, and very often these molecular orbitals (MO's) could be constructed by means of linear combination of atomic orbitals (LCAO) [33]. This idea forms the basis for the MO-LCAO approach, and it was first shown by Roothaan in the late 1940's and early 50's how this method can be formulated in a self-consistent-field way [34]. In this connection, a great deal of work had to be devoted to the problem of calculating one-, two, three-, and four-center *molecular integrals* involving atomic orbitals, e.g. Slater-type orbitals (STO's). With the development of the modern electronic computers, this problem could be partly circumvented by following Boys' [35] suggestion of using a network of Gaussian orbitals.

Let us assume that $\bar{\Phi} = \{\bar{\Phi}_1, \bar{\Phi}_2, \bar{\Phi}_3, \ldots, \bar{\Phi}_{M'}\}$ is a set of linearly independent one-electron functions with $M' \geq N$, having a metric matrix $\bar{\Delta} = \langle\bar{\Phi} | \bar{\Phi}\rangle$ and that the Hartree-Fock functions $\psi = \{\psi_1, \psi_2, \psi_3, \ldots\psi_N\}$ are formed by linear combinations of these basis functions, so that $\psi = \bar{\Phi}\,\mathbf{c}$, where $\mathbf{c}$ is a rectangular coefficent matrix of order $M' \times N$. If the set ψ is orthonormal, one gets directly the condition $\langle\psi|\psi\rangle = \langle\bar{\Phi}\mathbf{c}|\bar{\Phi}\,\mathbf{c}\rangle = \mathbf{c}^\dagger\bar{\Delta}\mathbf{c} = \mathbf{1}$. In such a case, the Fock-Dirac operator takes the form $\rho = |\psi\rangle\langle\psi| = |\bar{\Phi}\mathbf{c}\rangle\langle\bar{\Phi}\mathbf{c}| = |\bar{\Phi}\rangle\,\mathbf{c}\mathbf{c}^\dagger\,\langle\bar{\Phi}| = |\bar{\Phi}\rangle\mathbf{R}\langle\bar{\Phi}|$, where $\mathbf{R} = \mathbf{c}\mathbf{c}^\dagger$ is closely associated with the so-called charge- and bond-order matrix originally introduced by Coulson and Longuet-Higgins [36] Since the matrix $\mathbf{R}$ of order $M' \times M'$ has the properties $\mathbf{R}\,\bar{\Delta}\,\mathbf{R} = \mathbf{R}$, $\mathbf{R}^\dagger = \mathbf{R}$, $\mathrm{tr}\,(\mathbf{R}\bar{\Delta}) = N$, it reflects the properties of the projector ρ as expressed in (1.5). The more detailed properties of the matrix $\mathbf{R}$ and its importance for the SCF-procedure have been studied elsewhere [37].

Let us now study exactly the same approach when we like to separate the space and spin functions, and let us assume that $\Phi = \{\Phi_1, \Phi_2, \Phi_3, \ldots\Phi_M\}$ is a linearly independent *orbital basis* of order M with the metric matrix $\Delta = \langle\Phi | \Phi\rangle$. In such a case, one can construct a spin-orbital basis of the form $\bar{\Phi} = \{\Phi\alpha, \Phi\beta\}$ having $M' = 2M$ and the metric matrix

$$\bar{\Delta} = \begin{Bmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Delta \end{Bmatrix}. \qquad (2.22)$$

Let us express the orbital functions $\mathbf{u}$ and $\mathbf{v}$ in (2.5) in terms of the orbital basis Φ in the form:

$$\mathbf{u} = \Phi\,\mathbf{a}, \qquad \mathbf{v} = \Phi\,\mathbf{b},, \qquad (2.23)$$

where $\mathbf{a}$ and $\mathbf{b}$ are row vectors of order M. For the sake of simplicity, we will choose the orbital basis Φ *real*, whereas the coefficient vectors $\mathbf{a}$ and $\mathbf{b}$ may be *complex*. Hence we have

$$\psi = \mathbf{u}\,\alpha + \mathbf{v}\,\beta = \Phi\,\mathbf{a}\;\alpha + \Phi\,\mathbf{b}\,\beta = \bar{\Phi}\,\mathbf{c}, \qquad (2.24)$$

where

$$\mathbf{c} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \tag{2.25}$$

is a column vector of order M' = 2M, and further

$$\begin{aligned} \langle\psi|\psi\rangle &= \langle \Phi\,\mathbf{a}\;\alpha + \Phi\,\mathbf{b}\;\beta | \Phi\,\mathbf{a}\;\alpha + \Phi\,\mathbf{b}\;\beta\rangle = \\ &\mathbf{a}^\dagger\langle\Phi|\Phi\rangle\mathbf{a} + \mathbf{b}^\dagger\langle\Phi|\Phi\rangle\mathbf{b} = \\ &= \mathbf{a}^\dagger\Delta\,\mathbf{a} + \mathbf{b}^\dagger\Delta\,\mathbf{b} = \mathbf{c}^\dagger\,\bar{\Delta}\,\mathbf{c}, \end{aligned} \tag{2.26}$$

Recalling the notation $\mathbf{d} = \langle\psi|\psi\rangle^{-1}$, one obtains according to (2.8) for the Fock-Dirac operator

$$\begin{aligned} \rho &= |\psi\rangle\langle\psi|\psi\rangle^{-1}\langle\psi| = |\psi\rangle\,\mathbf{d}\,\langle\psi| = \\ &|\Phi\,\mathbf{a}\;\alpha + \Phi\,\mathbf{b}\;\beta\rangle\,\mathbf{d}\,\langle\Phi\,\mathbf{a}\;\alpha + \Phi\,\mathbf{b}\;\beta| = \\ &= \rho^{++}\,\alpha\alpha + \rho^{+-}\,\alpha\beta + \rho^{-+}\,\beta\alpha + \rho^{--}\,\beta\beta\,, \end{aligned} \tag{2.27}$$

where

$$\begin{aligned} \rho^{++} &= |\Phi\rangle\,\mathbf{a}\,\mathbf{d}\,\mathbf{a}^\dagger\langle\Phi| = |\Phi\rangle\,\mathbf{R}^{aa}\langle\Phi|\,, \\ \rho^{+-} &= |\Phi\rangle\,\mathbf{a}\,\mathbf{d}\,\mathbf{b}^\dagger\langle\Phi| = |\Phi\rangle\,\mathbf{R}^{ab}\langle\Phi|\,, \\ \rho^{-+} &= |\Phi\rangle\,\mathbf{b}\,\mathbf{d}\,\mathbf{a}^\dagger\langle\Phi| = |\Phi\rangle\,\mathbf{R}^{ba}\langle\Phi|\,, \\ \rho^{--} &= |\Phi\rangle\,\mathbf{b}\,\mathbf{d}\,\mathbf{b}^\dagger\langle\Phi| = |\Phi\rangle\,\mathbf{R}^{bb}\langle\Phi|\,, \end{aligned} \tag{2.28}$$

Here the quantities

$$\mathbf{R}^{aa} = \mathbf{a}\,\mathbf{d}\,\mathbf{a}^\dagger\,,\quad \mathbf{R}^{ab} = \mathbf{a}\,\mathbf{d}\,\mathbf{b}^\dagger\,,\quad \mathbf{R}^{ba} = \mathbf{b}\,\mathbf{d}\,\mathbf{a}^\dagger\,,\quad \mathbf{R}^{bb} = \mathbf{b}\,\mathbf{d}\,\mathbf{b}^\dagger\,, \tag{2.29}$$

are matrices of order M x M, which may be combined into a single matrix $\mathbf{R} = \mathbf{c}\,\mathbf{d}\,\mathbf{c}^\dagger$ of order 2M x 2M in agreement with the general theory. In the case when the Hartree-Fock functions ψ are orthonormal, one has of course $\mathbf{d} = \mathbf{1}$.

One can now substitute the extressions for $\mathbf{u}$ and $\mathbf{v}$ into the equations (2.21) provided than one observes that these equations are *exact* in the original Hartree-Fock scheme, whereas we are here using a truncated orbital basis of order M. Multiplying these relations to the left by $\langle\Phi(1)|$ and integrating over $\mathbf{r}_1$, one obtains instead the matrix equations

$$\begin{Bmatrix} \mathbf{h}+\mathbf{J}-\mathbf{K}^{++}; & -\mathbf{K}^{+-} \\ -\mathbf{K}^{-+}; & \mathbf{h}+\mathbf{J}-\mathbf{K}^{--} \end{Bmatrix} \begin{Bmatrix} \mathbf{a} \\ \mathbf{b} \end{Bmatrix} = \begin{Bmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Delta \end{Bmatrix} \begin{Bmatrix} \mathbf{a} \\ \mathbf{b} \end{Bmatrix} \varepsilon \,. \qquad (2.30)$$

In the standard way, it may be proven by means of the variation principle applied to the Fock-operator F(1) that these equations give the best approximations to the Hartree-Fock functions in the truncated orbital basis of order M. For the matrices involved, one gets the following expressions

$$\mathbf{h} = \langle\Phi(1)|H_1|\Phi(1)\rangle = \langle\Phi(1)|\mathbf{p}_1^2/2m - \Sigma_g\, e^2/r_{1g}|\Phi(1)\rangle, \quad (2.31)$$

$$\mathbf{J} = \langle\Phi(1)|J_1|\Phi(1)\rangle = e^2 \iint d\mathbf{r}_1 d\mathbf{r}_2\, \Phi(1) \frac{\rho^{++}(2,2)+\rho^{--}(2,2)}{r_{12}} \Phi(1), \quad (2.32)$$

$$\mathbf{K}^{\mu\nu} = \langle\Phi(1)|K_1^{\mu\nu}|\Phi(1)\rangle = e^2 \iint d\mathbf{r}_1 d\mathbf{r}_2\; \Phi(1) \frac{\rho^{\mu\nu}(1,2)}{r_{12}} \Phi(2), \quad (2.33)$$

where in the last line $\mu,\nu = +,-$. For the quatities $\rho^{\mu\nu}(1,2)$, one has further according to (2.28) and (2,29):

$$\rho^{\mu\nu}(1,2) = \Phi(1)\, \mathbf{R}^{\mu\nu}\, \Phi(2) = \Sigma_{\kappa\lambda}\, \Phi_\kappa(1) R^{\mu\nu}{}_{\kappa\lambda} \Phi_\lambda(2), \quad (2.34)$$

where on the right-hand sides one has now $\mu,\nu = a,b$. In the theory of molecular integrals, one is using the two notations

$$(\mu\kappa|\frac{e^2}{r_{12}}|\lambda\nu) = e^2 \int\int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\Phi_\mu(1)\Phi_\kappa(2)\Phi_\lambda(1)\Phi_\nu(2)}{r_{12}} = (\mu\lambda|\kappa\nu), \quad (2.35)$$

where the last one is the well-known "Mulliken notation", which refers to the two charge densities involved and has some obvious symmetry properties. For the matrix elements of the three fundamental matrices, one hence obtains

$$h_{kl} = \langle\Phi_k(1)|H_1|\Phi_l(1)\rangle = \langle\Phi_k(1)|\mathbf{p}_1^2/2m - \Sigma_g\, e^2/r_{1g}|\Phi_l(1)\rangle, \quad (2.36)$$

$$J_{kl} = \langle\Phi_k(1)|J_1|\Phi_l(1)\rangle = e^2 \int\int d\mathbf{r}_1 d\mathbf{r}_2\; \Phi_k(1) \frac{\rho^{++}(2,2)+\rho^{--}(2,2)}{r_{12}} \Phi_l(1) =$$

$$= \Sigma_{\kappa\lambda} (R^{aa}{}_{\kappa\lambda} + R^{bb}{}_{\kappa\lambda}) \, e^2 \int\int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\Phi_k(1)\Phi_\kappa(2)\Phi_\lambda(2)\Phi_l(1)}{r_{12}} =$$

$$= \Sigma_{\kappa\lambda} (R^{aa}{}_{\kappa\lambda} + R^{bb}{}_{\kappa\lambda}) \, (kl \,|\, \kappa\lambda), \qquad (2.37)$$

$$K^{\mu\nu}{}_{kl} = \langle \Phi_k(1) \,|\, K_1^{\mu\nu} \,|\, \Phi_l(1) \rangle =$$

$$= e^2 \int\int d\mathbf{r}_1 d\mathbf{r}_2 \; \Phi_k(1) \frac{\rho^{\mu\nu}(1,2)}{r_{12}} \Phi_l(2) =$$

$$= \Sigma_{\kappa\lambda} (R^{\mu\nu}{}_{\kappa\lambda} + R^{\mu\nu}{}_{\kappa\lambda}) \, e^2 \int\int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\Phi_k(1)\Phi_\kappa(1)\Phi_\lambda(2)\Phi_l(2)}{r_{12}} =$$

$$= \Sigma_{\kappa\lambda} (R^{\mu\nu}{}_{\kappa\lambda} + R^{\mu\nu}{}_{\kappa\lambda}) \, (k\kappa \,|\, \lambda l), \qquad (2.38)$$

where in the last relation one has $\mu,\nu = +,-$ or a, b, respectively. We note that, within the orbital subspace of order M, the equations (2.30) are identical with the original Hartree-Fock equations without any additional assumptions. We note further that, whereas the orbital basis Φ was assumed to be *real*, the vectors **a** and **b** may have *complex* character. From the point of view of interpretations, it is sometimes preferable to introduce - instead of the four components ρ^{++}, ρ^{+-}, ρ^{-+}, and ρ^{--}- the *number density matrix* $N(1,2)$ and the *spin density vector* $\mathbf{S}(1,2) = (S_1, S_2, S_3)$ defined through the relations

$$N(\mathbf{r}_1,\mathbf{r}_2) = \int \rho(\mathbf{r}_1,\zeta;,\mathbf{r}_2\zeta) \, d\zeta = \rho^{++}(\mathbf{r}_1,\mathbf{r}_2) + \rho^{--}(\mathbf{r}_1,\mathbf{r}_2), \qquad (2.39)$$

$$\mathbf{S}(\mathbf{r}_1,\mathbf{r}_2) = \int \mathbf{s} \, \rho(\mathbf{r}_1,\zeta;,\mathbf{r}_2\zeta) \, d\zeta \,, \qquad (2.40)$$

where the latter quantity has the three components:

$$\begin{aligned} S_1(1,2) &= [\, \rho^{+-}(1,2) + \rho^{-+}(1,2)]/2, \\ S_2(1,2) &= [\, \rho^{+-}(1,2) - \rho^{-+}(1,2)] \, i/2, \\ S_3(1,2) &= [\, \rho^{++}(1,2) - \rho^{--}(1,2)]/2, \end{aligned} \qquad (2.41)$$

see e.g. the first reference in [20] or [24].

The problem associated with the solution of the matrix equations (2.30) is to explore the energy surface $\langle H \rangle$ for the total Hamiltonian H, to study its local minima and extreme values, and - if possible - to find the absolute energy minimum connected with the Hartree-Fock method. Thanks to Fukutome's work [26], this problem has now been essentially

simplified, since one knows that every solution to the general Hartree-Fock equations corresponding to an energy minimum of <H> must belong to one of Fukutome's eight classes. The most restricted on is, of course, the RHF-solution, whereas the most general one is the one called "Torsional Spin Waves" (TSW).

Some Applications. - Looking at the literature, one finds that, in atomic and molecular physics, comparatively little interest has so far been devoted to the GHF-scheme, and the purpose of this paper is to try to focus the interests of some quantum chemists to this rather interesting problem of looking for the "absolute minimum" of the original Hartree-Fock method.

Since the general spin-orbitals (GSO's) of type (2.4) contain twice as many orbital functions as the one-electron functions used in the UHF-scheme, one could perhaps expect that, even for atomic and molecular systems, the GHF-method would give a lower variational energy than the UHF-method [38]. Applications to some two-electron systems - the helium-like ions and the hydrogen molecule - by Lunell [39] showed, however, that the GHF-scheme converged to exactly the same energy as the UHF-scheme. Physically this depends probably on the fact that a two-electron system has its lowest energy when the two electrons have opposite spins. As an introduction, we repeated these calculations by means of somewhat different computational tools - with starting points also in the complex plane - with the same result, but we still have not been able to give a simple mathematical proof that, for two-electron systems, the GHF- and the UHF-methods give the same result. This is hence still an open question.

Even if some interesting applications of the GHF-method had been found in solid-state theory [23,24], the applications to molecular systems were comparativlely few [40]. One major application to molecular systems had been worked out by Fukutome [40], and it was a study of the properties of the polyacetylene by means of the Pariser-Parr-Pople (PPP) approximation. It seemed hence desirable to make a molecular study based on *ab-initio* calculations to verify that one would get similar results and to get some experience in handling general Hartree-Fock orbitals of a complex nature, and for this purpose we started with some simple applications to atoms and to the BH molecule.

3. Applications of the General Hartree-Fock Theory to some Atoms and the BH Molecule

The programming of the GHF equations (2.21) in matrix form (2.30) was carried out in a straightforward manner: the one- and two-electron integrals were calculated by a slightly modified version of the well-known GAUSSIAN-70 package [41]. The SCF calculations were carried out by using rather strict convergence criteria: a numerical convergence of about 10^{-10} a.u. in the energy and 10^{-7} to 10^{-8} in each element of the matrix **R** have been required as the largest permitted changes in an SCF cycle. The achievement of this degree of convergence often required a very large number of iteration steps in the SCF procedure (sometimes several hundred or even one or two thousand), because the energy hypersurfaces as functions of the orbital coefficients turned out to be extremely flat. (This behaviour was partly due to the existence of some orientational arbitrariness in the problems studied, which will be discussed later, and in some cases due also to the occurrence of different solutions in very near vicinity of each other). No extrapolation methods were applied to speed up the convergence in order to avoid "jumps" from one branch of solutions to the other. Since we used a basis consisting of real Gaussians, we have carefully considered the case of complex LCAO coefficients, and it is hence remarkable that all the wave functions actually obtained in the calculations were expressible by means of real coefficients.

Some crucial aspects of studying the GHF wave functions are connected with the relationship between the GHF and UHF methods. First of all, it is evident that the RHF and UHF wave functions are particular solutions also to the GHF problem: In this case the components ρ^{+-} and ρ^{-+} of the Fock-Dirac density matrix are zero, and the GHF equations separate into two sets of equations for the orbitals of spins α and β, respectively. The system of equations obtained in this way is identical to that of the ordinary UHF scheme. We note that the two sets of equations are still coupled through the components ρ^{++} and ρ^{--}. The situation is in some way analogous to the case of the UHF equations for a closed-shell system, for which the RHF functions always provide a particular solution. Similarly to the RHF versus UHF case, the UHF (or RHF) solution can , in principle, represent either a true (local) minimum or a saddle point for the GHF problem.

There is another aspect which makes it somewhat difficult to find genuine GHF solutions. Since the conventional RHF and

UHF schemes are characterized by vanishing components ρ^{+-} and ρ^{-+}, it is obviously a *necessary* condition for the existence of the GHF solution that these components are non-vanishing: $\rho^{+-} = (\rho^{-+})\dagger \neq 0$. At the same time, it should be observed that this condition is far from *sufficient*, since the energy is independent of the axes of the spin quantization [39]. The basic Hamiltonian (1.1) does not contain the spin, and it is invariant under rotations of the underlying three-dimensional Euclidean space. When the coordinate space (x,y,z) is rotated, the spin functions undergo a unitary transformation:

$$\alpha' = \alpha\, u_{11} + \beta\, u_{21}, \qquad \beta' = \alpha\, u_{12} + \beta\, u_{22}, \qquad (3.1)$$

where the coefficients form a two-dimensional unitary matrix $\mathbf{u} = \{u_{kl}\}$. If one replaces the spin functions α and β in a set of N conventional UHF function ψ with α' and β', one obtains a set of new UHF function ψ' which look like GHF functions (1.7) and which have a Fock-Dirac density matrix with $\rho^{+-} = (\rho^{-+})\dagger \neq 0$, but which still have the same orbital energies and total energy. If one starts from the assumption that a *genuine GHF solution*, due to the richer possibilities for variation, should have a lower energy than the UHF solution, then one has to exclude the spin rotated UHF functions. In our calculations, we hence checked this possibility by investigating whether the matrix $\mathbf{R}^{ab}$ defined by (2.29) could be reduced to zero by a spin rotation of type (3.1). For a genuine GHF solution, such a reduction should, of course, not be possible.

The simplest system we considered was the H_2 molecule treated at the STO-6G and 4-31G levels; for this molecule, however, we could not obtain any GHF solution with an energy lower than the UHF one. In the minimum basis case, we have also explored the energy hypersurface in somewhat greater detail by introducing a dense mesh of the independent parameters defining the two-electron wave function and calculating its energy in all points of the mesh; even in this way no GHF energy lower than the UHF energy could be found. This result seems to indicate that in this case there is no genuine GHF solution, and that it is identical to the UHF solution. As we mentioned at the end of the previous section, this result for a two-electron system depends probably *physically* on the fact that the total energy is lowest when the spins α and β - or more generally α' and β' - have opposite spin directions. It remains to prove this statement also mathematically.

In studying atoms, no genuine GHF solutions could be obtained for the case of the He and Li atoms, either, in accordance with some previous results [39].

The simplest system for which a specific GHF solution has been obtained is the Be atom treated at the STO-6G level. Table 1 summarizes the energy values obtained for the $(1s)^2(2s)^2$ ground state of this atom by the RHF, UHF, and GHF, as well as those obtained for the triplet and singlet excited states. (The triplet state corresponded to $S_z = 1$ and was calculated by the UHF scheme, while the excited singlet describes an RHF configuration). For the UHF wave function a simultaneous breaking of both spatial and spin symmetries takes place, and the UHF energy is significantly lower than the RHF one. This UHF energy lowering (1.7×10^{-3} a.u.) can most likely be related to the existence of a low-lying triplet state and to the "quasi-degeneracy" between between the 2s and 2p orbitals reducing the value of the first singlet-singlet excitation energy. The GHF method gives an additional energy lowering of 1.2×10^{-4} a.u.; this relationship between the energy gains in the UHF and GHF methods can be considered rather typical also for the other cases we have considered.

Both the UHF and the GHF methods give a partial description of the angular correlation in the Be atom: there is a small admixture of the 2p basis orbitals in the SCF orbitals describing the 2s shell (and, to a less extent, the 1s shell). For beryllium treated in a minimal basis set, there is no possibility to describe any radial correlation. The GHF wave functions obtained for models with split valence shell (6-31G or 4-31G) basis sets, which we will discuss below, give a partial account of both radial (or left-right) and angular correlation. However, the effect of the latter seems to be crucial for the very existence of the genuine GHF solutions, similar to the specific (both spin- and symmetry-unrestriced) UHF solutions to which they are close. A comparison with Fukutome's scheme shows that these GHF functions are Torsional Spin Density Waves (TSDW).

The typical GHF picture may perhaps best be seen by considering the expansion of the pair of GHF 2s orbitals of the Be atom:

$$\psi_3 = -0.270\varphi(1s).\alpha + 0.996\varphi(2s).\alpha + 0.189\varphi(2p_y).\alpha + 0.189\varphi(2p_x).\beta, \qquad (3.2)$$

$$\psi_4 = -0.270\varphi(1s).\beta + 0.996\varphi(2s).\beta - 0.189\varphi(2p_y).\beta + 0.189\varphi(2p_x).\alpha,$$

These equations indicate that both 2s orbitals have admixtures of two different p-orbitals, one with α spin and one with β spin, with equal absolute values of the coefficients but with a characteristic difference in the phases for the two orbitals. The equal absolute values of the coefficients imply that such pairs of GHF orbitals of TSDW type have identical orbital energies.

The next system under consideration was the carbon atom. Table 2 shows the energies obtained for the carbon atom by using 4-31G and 6-31G basis sets. The overall behaviour of the solutions is similar to that in the beryllium case. For carbon, however, the true ground state is a triplet in accordance with Hund's rule. The GHF energy obtained is significantly lower than the singlet RHF one, but it is higher than the triplet UHF value. This means that the GHF wave function obtained represents a mixture of singlet and triplet states, which leads to a stationary value of the energy but not to its absolute minimum. In such a case, it is probably not meaningful to discuss to what extent electron correlation is accounted for by the GHF method. It does not seem to be excluded that another GHF solution with an energy lower than the triplet UHF may exist, but we made no special search for such a solution, since we wanted to go on to the molecular case.

Several calculations for the BH molecule at the 4-31G level were also performed. This molecule is a very interesting species: it is a closed-shell σ-system which is, at the same time, paramagnetic. In the RHF or conventional (spin-unrestricted) UHF approaches, all the σ-orbitals are empty; its characteristic magnetic behaviour may then tentatively be connected with the low excitation energies needed to transfer an electron to the low-lying excited σ-levels.

For the BH molecule *six different solutions* of the GHF equations were obtained: one RHF solution, three different UHF solutions, and two different genuine GHF solutions of TSDW type. A part of these solutions exists only at some values of the intermolecular distance; there is, however, a narrow interval for which all six solutions were obtained simultaneously.

Fig. 1a shows an overall view of the potential curves of the BH molecule corresponding to the different SCF solutions. The RHF solution exists, in principle, at all intermolecular distances R, although - for $R > 3.5$ Å - the conventional SCF procedure does not converge. (This is quite typical for similar systems). At large distances, the RHF curve shows the well-known incorrect

Table 1.

Energies of the Be atom calculated by using a STO-6G basis set.

Wave function	Energy in a.u.
RHF	-14.503 361
UHF(S_z=0)	-14.505 074
GHF	-14.505 190
Triplet (UHF,S_z=1)	-14.442 082
Singlet $(1s)^2(2p)^2$	-14.172 547

Table 2.

Energies of the carbon atom calculated by using 4-31G and 6-31G basis sets.

Wave function	Energies in a.u. 4-31G	6-31G
RHF	-37.544 557	-37.588 204
UHF (S_z=0)	-37.604 05	-37.647 030
GHF	-37.612 63	-37.655 524
Triplet (UHF, S_z=1)	-37.635 05	-37.677 837

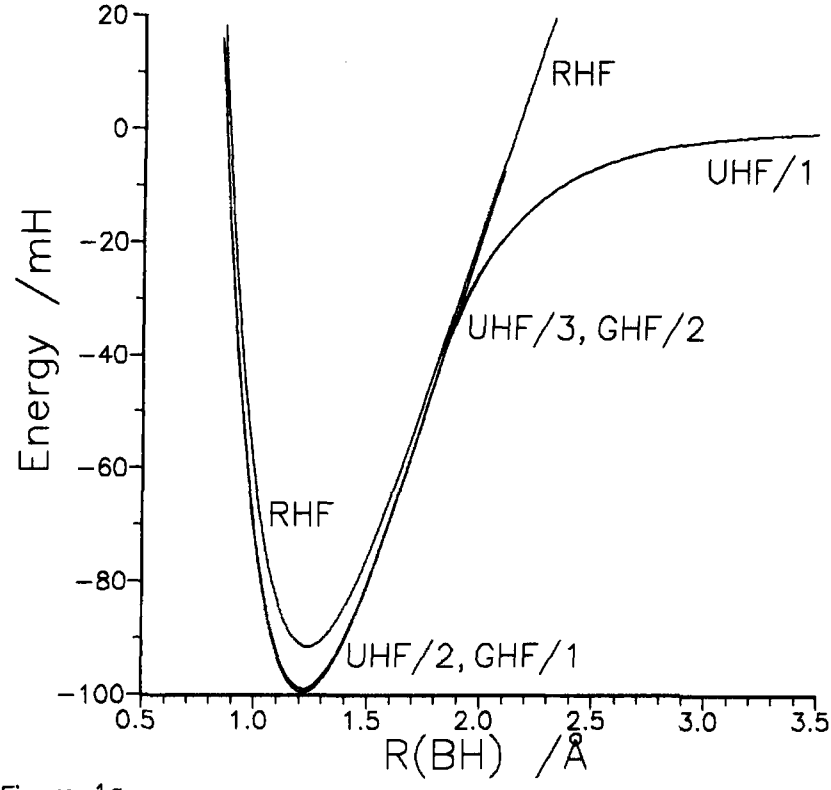


Figure 1a

dissociation behaviour [42] connected with the large weight of the ionic configurations in the RHF wave functions built up of doubly filled orbitals. This wrong asymptotic behaviour is corrected by the usual UHF wave function (denoted here as UHF/1). At R ≈ 1.8 Å, the UHF/1 potential curve smoothly departs from the RHF one and tends to a correct limiting value at $R \to \infty$, giving a basically adequate description of the main features of the left-right correlation. However, as discussed in [43], the UHF potential curve makes a short turn after departing from the RHF one and approaches the (correct) limiting value too quickly. This is connected with the sudden appearance of the specific UHF solution differing from the RHF one, and indicates that the UHF wave functions do not have enough flexibility to describe properly the binding effects and to reduce effectively the weight of the ionic terms. Therefore, there is a too quick transition between the limiting cases for which essentially one of these factors is taken into account. (The RHF method is better for describing the binding effects, and the usual UHF solution appears only when they are overweighted by the left-right correlation.) As discussed in [43], this defect cannot be eliminated by using single determinant wave functions, while it does not occur in the projected (extended) Hartree-Fock method. (In the present case, besides the usual RHF and UHF wave fucntions, there are several other single determinant SCF wave functions of the UHF and GHF type. Neither of them solves, however, this problem of the too fast transition between the bonding and asymptotic regions.) In the bifurcation point, the UHF/1 potential curve merges continously with the RHF one and has also a first derivative, coinciding with that of the RHF wave functions, but there is a discontinuity in its second derivative [43].

Analogously to all similar UHF wave functions describing dissociation of a closed shell molecule into two odd-electron fragments, the UHF/1 wave functions is characterized by a pair of molecular orbitals of opposite spins, which are more and more localized on the boron and on the hydrogen, respectively, as the internuclear distance R increases. (In the limit $R \to \infty$, one of these orbitals becomes a pure boron, another a pure hydrogen atomic orbital).

At the equilibrium internuclear distance, there are two solutions with energies lower than the RHF one (Fig. 1b). The curve denoted UHF/2 corresponds to a spin- and spatially-unrestricted single determinant wave function, in which the σ-orbitals contain some admixture of the π-type 2p boron orbitals, which are empty in both the RHF and UHF/1 wave functions.

The GHF/1 solution differs from the UHF/2 one by a small admixture of π-orbitals of opposite spin. This is a genuine GHF solution which cannot be reduced to a UHF one by any rotation of the spin axes.

All the three solutions existing in the interval around the equilibrium internuclear distance give the position of the minimum on the potential curve at about 1.22 - 1.23 Å, which is in good agreement with the experimental value 1.232 Å [44]. Neither of the UHF/2 or GHF/1 wave functions describe any left-right correlation: in both cases there are three pairs of molecular orbitals characterized by degenerate orbital energies and strictly equal coefficients of the boron σ-orbitals as well of the hydrogen AO's. These orbitals differ only in the admixtures of the π-orbitals. in a manner analogous to the beryllium example described in (3.2). Accordingly the UHF/2 and GHF/1 potential curves exhibit a wrong dissociation behaviour, quite analogous to that of the RHF curve.

Due obviously to the increasing importance of the left-right correlation, the UHF/2 and GHF/1 solutions become *numerically* unstable at larger internuclear distances. At the distance R = 2.05 Å, the SCF procedure started from the orbitals obtained at R = 2.00 Å give first convergence to a UHF/2 or GHF/1 wave function, respectively. (Convergence up to ~10^{-5} in the R-matrix elements and ~10^{-8} a.u. in the energy has been obtained). Then - under the influence of the rounding-off errors - the calculation departed slowly from the domain of the wave functions of type UHF/2 and GHF/1 and gave a (rather slow) convergence to wave functions of a new type and significantly lower energy. These wave functions are denoted by UHF/3 and GHF/2; they can be considered as combinations of the UHF/2 or GHF/1 wave functions describing the angular (σ - π) correlation and the UHF/1 function describing left-right correlation. Accordingly, as Fig. 1c shows, the UHF/3 and GHF/2 potential curves span an arc between the UHF/2 or GHF/1 curve and the UHF/1.

It should be observed that, above R ~ 2.1 Å, only the UHF/1 solution could be obtained. Similarly, below R = 1.82 Å, the numerical effects cause a "jump" from the UHF/1 branch to the UHF/2 one in the course of the SCF procedure, since UHF/2 has a lower energy at that distance. This means that there is an interval of a length of about 0.2 Å, in which all the six different SCF solutions exist simultaneously. The situation is best characterized by Figs. 2a and 2b, in which the *energy lowering with respect to the RHF method* given by the different UHF and

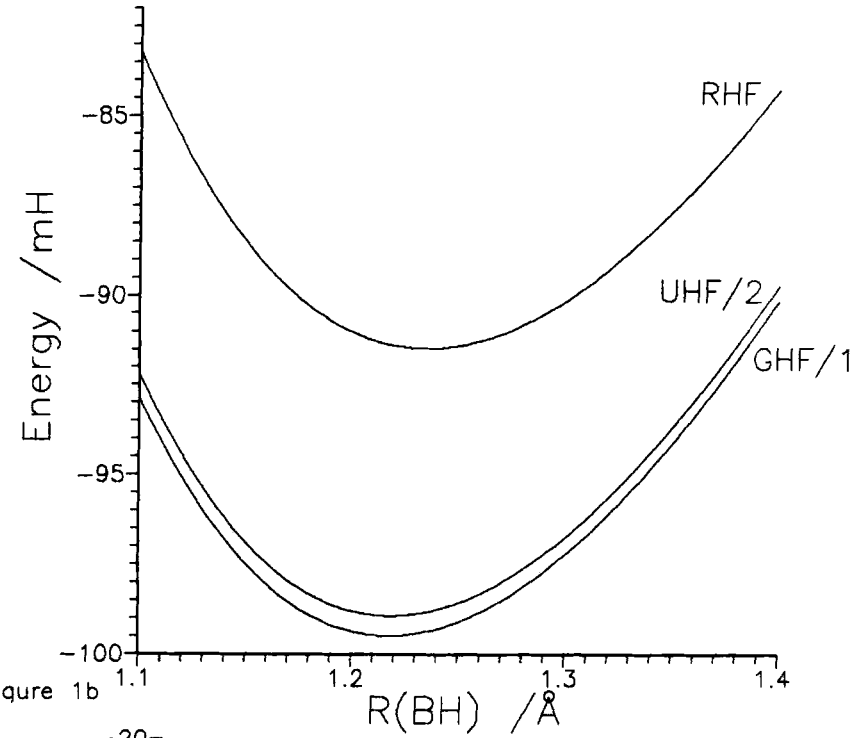


Figure 1b

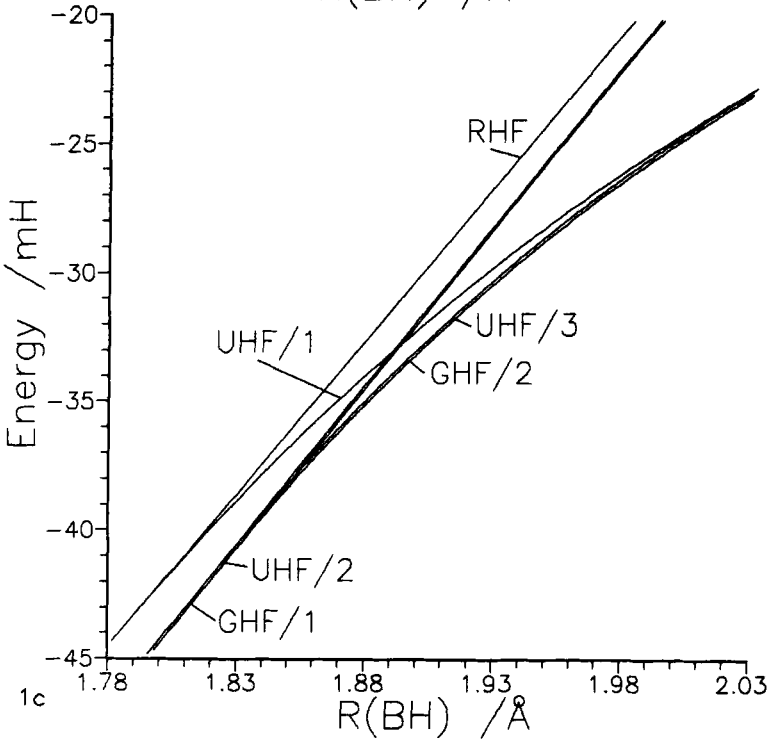


Figure 1c

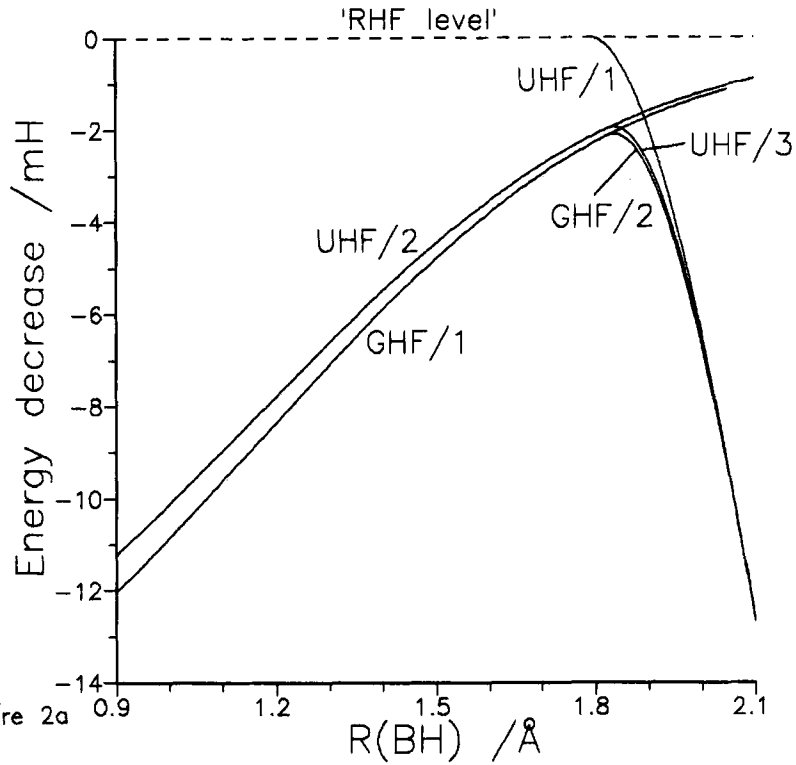


Figure 2a

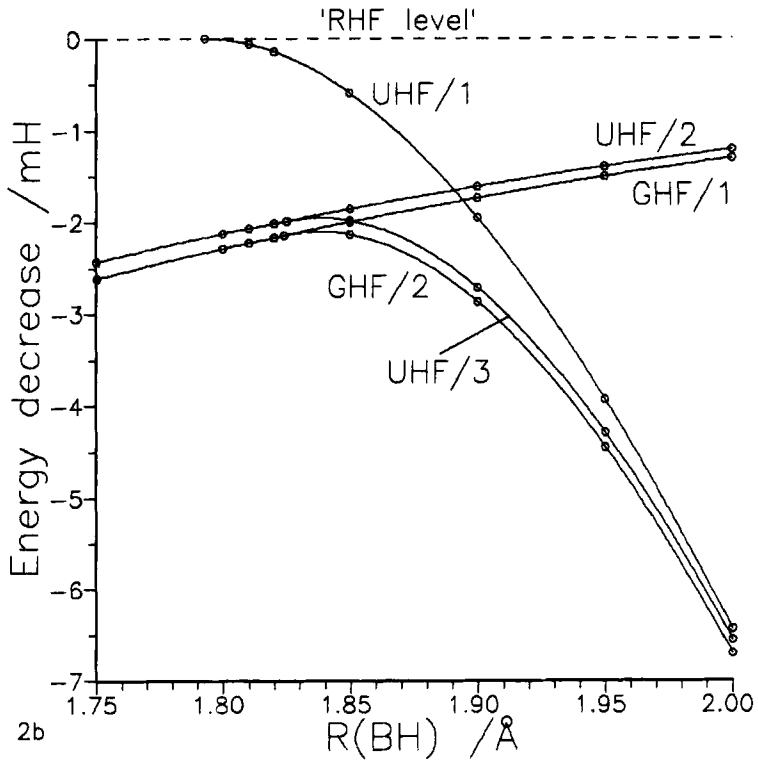


Figure 2b

GHF methods is presented. These curves show rather clearly the characteristics of the different particular solutions obtained.

At the smaller internuclear distances, the UHF and GHF methods give a partical account of the σ - π angular correlation (the UHF/2 and GHF/1 curves), while at the larger distances the left-right correlation described by the UHF/1 solution is the dominating effect. There is a narrow interval where these two effects are of a comparable magnitude: here the "usual" UHF/1 curve crosses the UHF/2 and GHF/1 ones. (This is also the interval in which the short turn on the UHF/1 curve, discussed above, is just started: the importance of the left-right correlation is quickly increasing). The most interesting result is, in our opinion, that there is not only a competition between the effects of angular and left-right correlation, but also that their combination is possible, giving rise to the specific solutions UHF/3 and GHF/2 exhibiting a further energy lowering. This stresses again the mathematical complexity of the Hartree-Fock problem due to the pronounced non-linear character of the general Hartree-Fock equations.

4. A Study of the Hessians in the Hartree-Fock Scheme

The multiplicity of the Hartree-Fock solutions obtained for the BH molecule motivated us to perform a special study to clarify which ones of these solutions correspond to true (local) minima, and which ones correspond to saddle points on the energy hypersurface, and to determine the bifurcation points in which new types of solutions appear as exactly as possible. For that reason we have investigated the Hessians for the RHF, UHF/2 and GHF/1 wave functions discussed above. As is well-known [19,20], the Hessian is defined as a matrix $\mathbf{H} = \{H_{ij}\}$ with the elements

$$H_{ij} = \partial^2 E / \partial c_i \, \partial c_j, \qquad (4.1)$$

where E is the total energy and the quantities c_i are the independent parameters defining the wave function Ψ under consideration. If all the eigenvalues of $\mathbf{H}$ are positive, the stationary point considered is a true minimum; if one or more of the eigenvalues are negative, it is a saddle-point. A zero eigenvalue indicates that there exists a parameter, of which the energy is independent at least up to the second order.

In the case of a single determinant wave function, the parameters c_i above could be, in principle, identified with the

LCAO coefficients of the SCF orbitals involved. However, the invariance of the determinantal wave functions under non-singular linear transformations of the N basic one-electron functions indicates that the number of really independent parameters is less than that of the LCAO coefficients. Instead of using them, we may utilize a theorem [45] which in the present case can be formulated as follows: the N not necessarily normalized or orthogonalized one-electron functions ψ_i in an arbitrary determinantal wave function may be presented by the expression:

$$\psi_i = \psi_i^{occ} + \sum_p \psi_p^{virt} \eta_{pi}, \qquad (4.2)$$

where ψ_i^{occ} and ψ_p^{virt} are the occupied and virtual SCF orbitals corresponding to the solution under study, the summation over p goes from 1 to (M-N), and M is the dimension of the spin-orbital basis used in the calculations. In this case, the expansion coefficients η_{pi} play the role of the independent parameters c_i used in (4.1).

Even if one starts from SCF estimates of a complex nature, only *real* SCF wave functions have been obtained as a result of the calculations, and this means that both the occupied and virtual SCF orbitals form a real orthonormalized set, spanning the whole subspace of the LCAO basis orbitals, as noted previously. This implies also that all the matrix elements of the Hamiltonian between determinantal wave functions built up from these SCF orbitals are necessarily real. Utilizing this fact, and following the derivation given in [20], the Hessian can be given as

$$\mathbf{H} = \begin{bmatrix} \mathbf{A+B} & \mathbf{0} \\ \mathbf{0} & \mathbf{A-B} \end{bmatrix}, \qquad (4.3)$$

where the matrices **A** and **B** are given by the relations

$$A_{kl}^{pr} = \langle \psi_k^p | H - 1 | \psi_l^r \rangle, \quad B_{kl}^{pr} = \langle \psi_{kl}^{pr} | H - 1 | \psi_0 \rangle, \qquad (4.4)$$

and the wave functions ψ_k^p and ψ_{kl}^{pr} are determinants which are singly and doubly excited with respect to the SCF determinant Ψ_0:

$$\psi_k^p = \partial\Psi/\partial\eta_{pk} = \Psi_1(\psi_k^{occ} \rightarrow \psi_p^{virt}), \qquad (4.5)$$

$$\psi_{kl}^{pr} = \partial^2\Psi/\partial\eta_{pk}\partial\eta_{rl} = \Psi_2(\psi_k^{occ} \rightarrow \psi_p^{virt}, \psi_l^{occ} \rightarrow \psi_r^{virt}). \qquad (4.6)$$

Here Ψ is the determinant wave function built up from N spin-orbitals of type (4.3), and the derivatives are taken in the point $\Psi = \Psi_0$, i.e. for all $\eta_{pk} = 0$.

The eigenvalue problem of the Hessian matrix **H** factorizes according to (4.3) into the eigenvalue problems for **A+B** and **A-B**, the former associated with the stability of Ψ_0 under pure real variations and the latter under pure imaginary ones [20].

The derivation of the explicit formulas and programming of the elements of the matrices $\mathbf{A} \pm \mathbf{B}$ and their diagonalization was straightforward. The first interesting result obtained was the observation that the lowest eigenvalue of the Hessian in the case of the free Lithium atom is zero. This can easily be understood if one takes into account the fact that the energy of the Li atom in its ground state $(1s)^2(2s)^1$ is obviously independent of whether the unpaired electron has α spin or β spin - or a spin quantized along an arbitrary axis in the space.

Similarly, there is at least one zero eigenvalue for the GHF/1 and UHF/2 Hessians studied for the BH molecule, but not for those corresponding to the RHF solution.This is due to the fact that the UHF/2 and GHF\1 wave functions do not have the C_∞ symmetry of the molecule, but only a (quasi)symmetry of the C_{2v} type. Therefore, a rotation of the π-contamination (which is added to the σ-orbitals) around the internuclear axis does not change the total energy.

Figs. 3a and 3b show the lowest eigenvalues of the Hessian for the RHF wave function as functions of the internuclear distance. One can see that there is at least one negative eigenvalue at every distance, i.e. the RHF function has a saddle point character everywhere. (The existence of the UHF and GHF solutions with energies lower than the RHF would not necessarily imply that conclusion, since the RHF solution could have, in principle, also a true *local* minimum.) One of the eigenvalues of the Hessian sharply decreases with increasing internuclear distance and at R = 1.793 Å it crosses the zero level: this is the bifurcation point in which the specific UHF/1 solution is appearing.

The lowest eigenvalues of the Hessian for the UHF/2 wave function are displayed in Figs. 4a and 4b. There is everywhere a negative eigenvalue which can be connected with the existence of the GHF/1 solution with a lower energy. The overall picture is

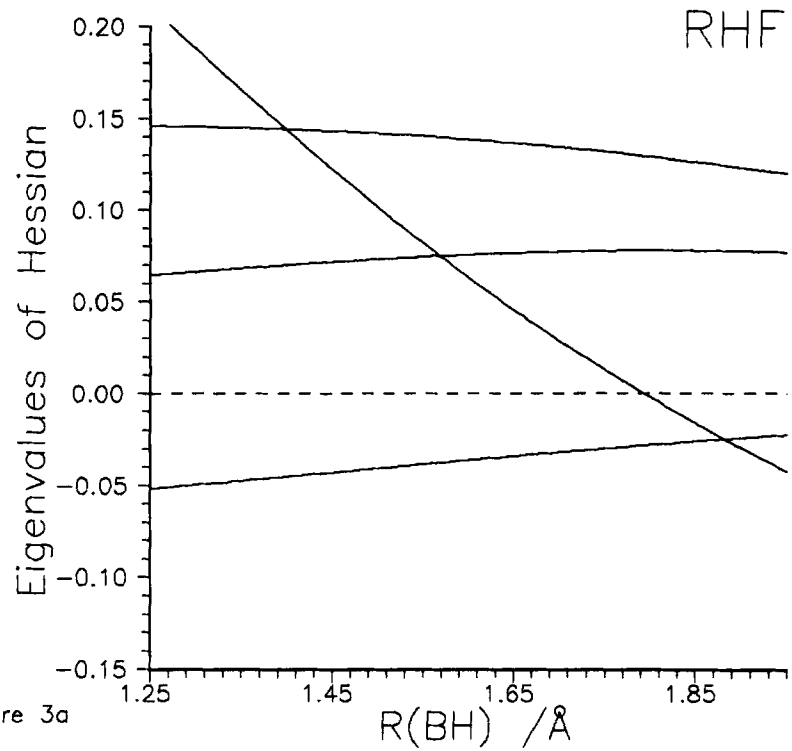


Figure 3a

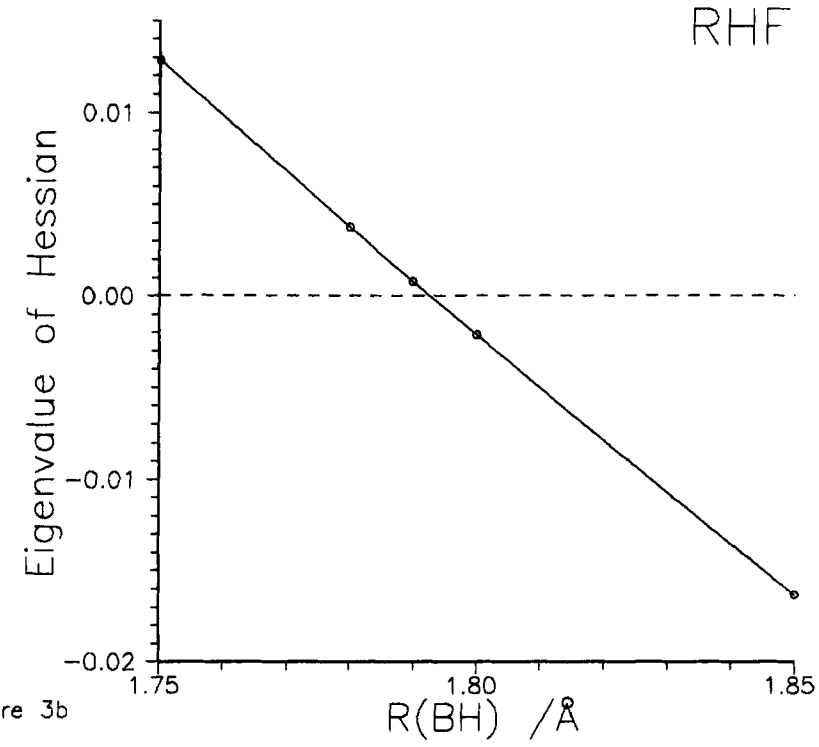


Figure 3b

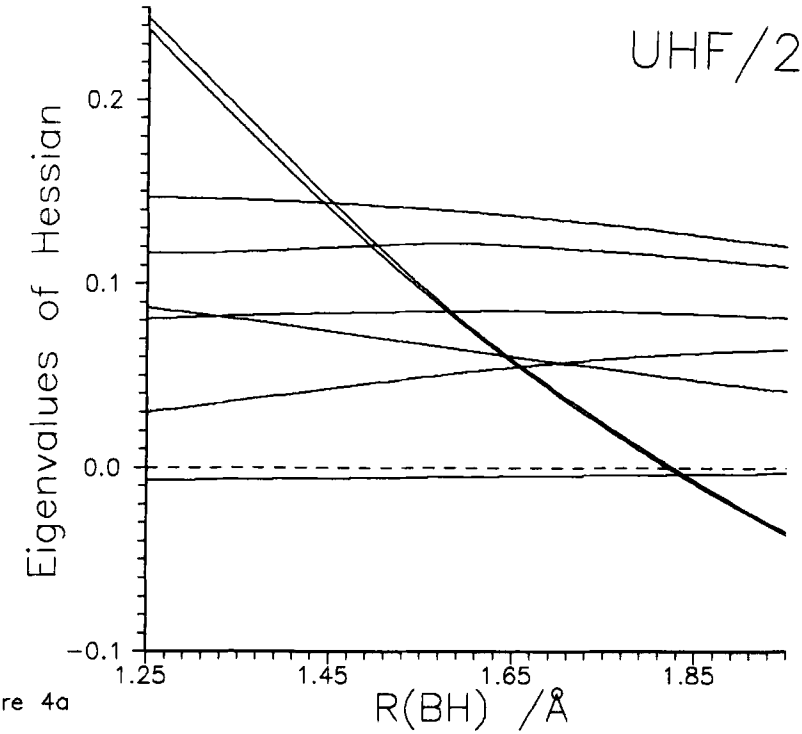


Figure 4a

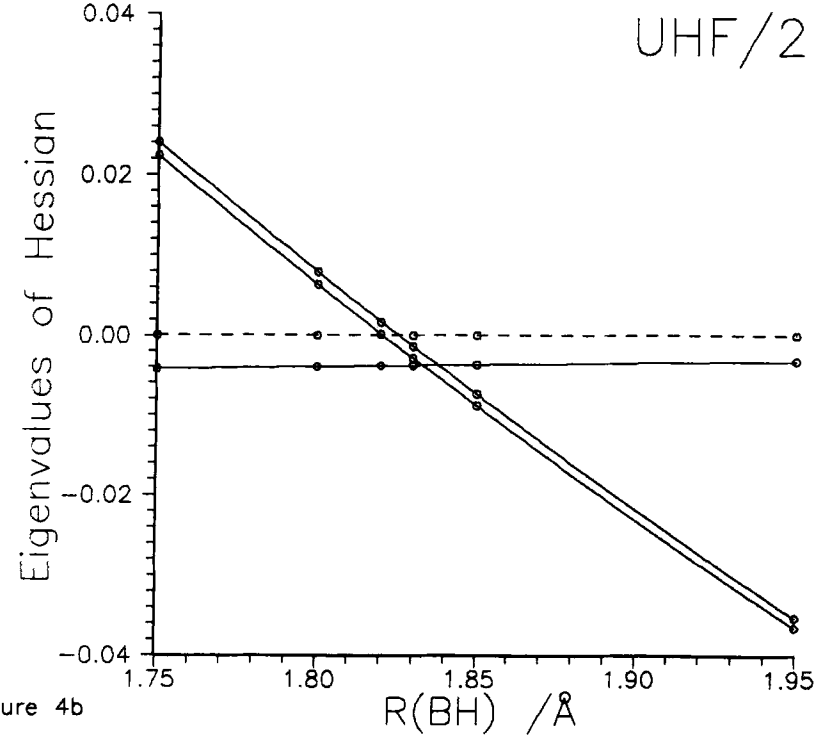


Figure 4b

more complicated than in the RHF case, due to the zero eigenvalue present at all internuclear distances (see above) and to the fact that a pair of close eigenvalues crosses the zero level at R ≈1.82 Å. They are related to the appearance of a pair of new solution UHF/3 and GHF/2 around this point.

The lowest eigenvalues of the Hessian for the GHF/1 solution are depicted in Figs. 5a and 5b. The picture is now somewhat simpler: there is a zero eigenvalue at all distances, connected with the rotational arbitrariness mentioned above, but there is no negative eigenvalue until R ≈1.824 Å, which is the point where the GHF/2 solution starts to exist.

In the case of the RHF and UHF/2 wave functions, the matrix **A-B** does not have any eigenvalues which are not present also for the matrix **A+B** ; only the degrees of degeneracy may be different. However, for the GHF/1 Hessian no such coincidence of eigenvalues is present. Moreover, the eigenvalues of the matrix **A-B** crosses the zero level at a point about a thousand of an Å prior to the eigenvalues of the matrix **A+B** . Based on this observation, we have tried to find a further, *substantially complex*, solution to the GHF equations, utilizing the corresponding eigenvector of the Hessian for this purpose. This attempt was unsuccessful with the computational tools we had available; nevertheless, one cannot exclude the existence of such a seventh solution of the GHF equations. We may suppose, however, that even if such a solution does exist mathematically, one may have to go to much higher numerical accuracy to be able to distinguish it from the GHF/2 solution.

Our little study shows as a side result that, in the solution of the SCF problem at any level, it is essential to understand in much greater detail the problem of*the connection between the starting point of the SCF procedure and the final SCF solution* than we do today. In our opinion, it would hence be worthwhile to devote much more research to this problem - both mathematically and numerically.

5. Concluding Remarks

In this paper we have studied the original Hartree-Fock method - here referred as the GHF-scheme - for some simple atomic and molecular systems by using general spin-orbitals (1.7) of a complex character, and we have found that there exist GHF-solutions of TSDW type which give lower energies <H> than the RHF- and UHF-schemes, but also that quite a few SCF solutions of different type may exist simultaneously. The purpose of the

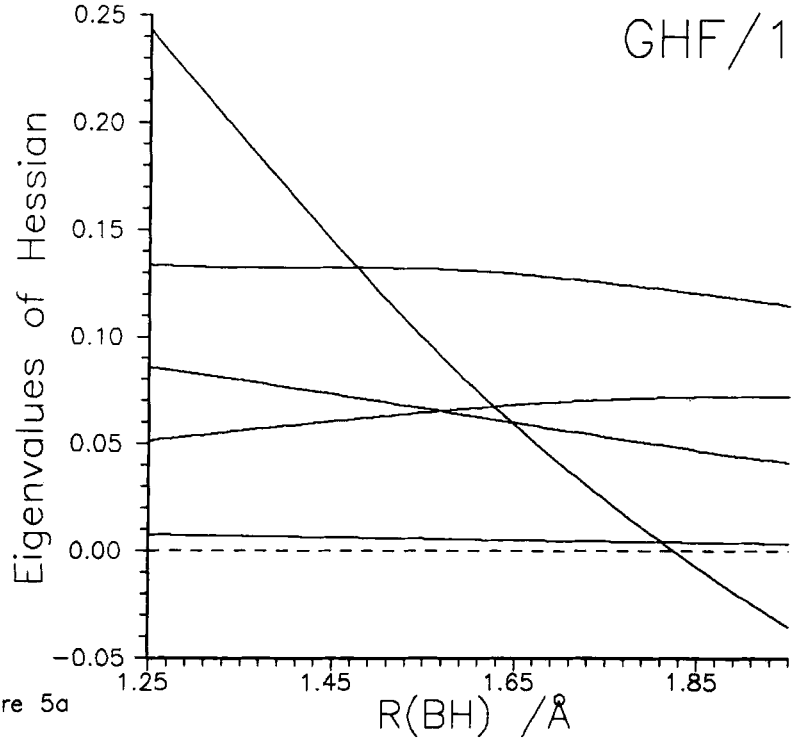


Figure 5a

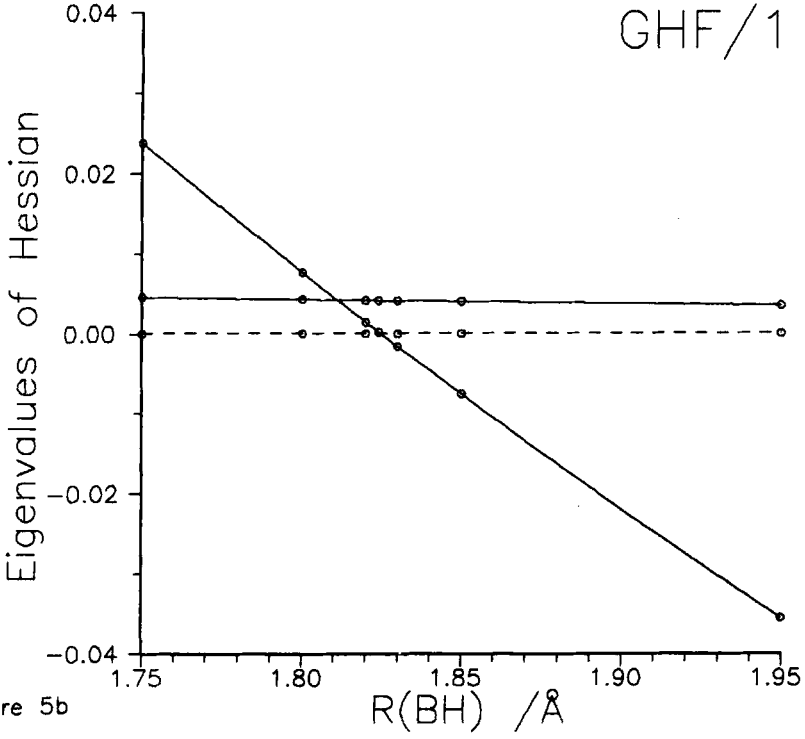


Figure 5b

study is, of course, to find the *absolute minimum* <H> for the energy on the hypersurface associated with the SCF procedure. At the same time it should be observed that, whenever such an absolute minimum is found, the associated Slater determinant D is usually a *mixture* of different symmetry types. It is true that by means of the symmetry projectors Q_k forming a resolution of the identity $1 = \Sigma_k Q_k$, one may carry out a uniqe component analys of the Slater determinant D:

$$D = 1.D = (\Sigma_k Q_k)D = \Sigma_k Q_k D = \Sigma_k D_k , \qquad (5.1)$$

where the component $D_k = Q_k D$ has the correct symmetry properties. It is also clear that this component does not have an optimal energy $\langle H \rangle_k$, since one has varied before the projection, and that - if one want to preserve the correct symmetry property - one ought to take the projection before the variation. In this way one is naturally lead to the Projected Hartree-Fock (PHF) scheme [46], which preserves many of the fundamental features of the independent-particle-model. So far, the PHF-scheme has in the DODS- and AMO-methods [47] been developed essentially as an extension of the UHF-scheme, but it is obvious that one could get an even better treatment of the correlation problem, if one could take the projection of a Slater determinant built up from *general spin-orbitals* . It has been shown [48] that, even for two-electron systems, the PHF scheme gives a lower energy when one is using general spin-orbitals (GSO's) and that the same is true also for the lithium atom [49] with N = 3. So far it has not been easy to carry out an extension of the PHF-scheme using complex GSO's to many-electron systems. In such a case, it is certainly important to learn how to handle determinants built up from complex general spin-orbitals, and it is hence hoped that the study in this paper will be of value not only as an exposé of the GHF method but also as a first step towards an extension of the PHF-sceme.

Acknowledgements.- The authors are indebted to Professors Jean-Louis Calais and Sten Lunell of the Uppsala Quantum Chemistry Group for critically reading the manuscript and for valuable suggestions. They are also grateful to Dr. Jozef Noga for his kind help in adapting the GAUSSIAN-70 program to the SUN system at the Florida Quantum Theory Project.

References

1. N. Bohr: Proc. Phys. Soc. (London) **35**, 296 (1923).
2. D. R. Hartree: Proc. Cambridge Phil. Soc. **24**, 89 (1928).
3. D. R. Hartree: I Waller:
4. J. C. Slater: Phys. Rev. **34**, 1293 (1929).
5. M. Born and J. R. Oppenheimer: Ann. der Physik (Liepzig) **84**, 457 (1927).
6. J. C. Slater: Phys. Rev. **34**, 1293 (1929); **35**, 210 (1930).
7. V. Fock: Z. Physik **61**, 126 (1930).
8. P.A.M. Dirac, Proc. Cambridge Phil. Soc. **26**, 376 (1930); **27**, 240 (1931).
9. T. A. Koopmans: Physica **1**, 104 (1933).
10 D. R. Hartree: Mem. Proc. Lit. Phil. Soc. Manchester **77**, 91 (1933).
11. V. Fock and M. Petrashen: Phys. Zs. Sow. **6**, 368 (1934); **8**, 547 (1935).
12. M. Delbrück, Proc. Roy.Soc. (London) **A129**, 686 (1930).
13. J. C. Slater: Phys. Rev. **81**, 385 (1951); **82**, 538 91951); Revs. Mod. Phys. **25**, 199 (1953).
14. R. K. Nesbet, Proc. Roy.Soc. (London) **A230**, 312 (1955); G. W. Pratt, Jr.,Phys. Rev. **102**, 1303 (1956); J.H.Wood and G.W. Pratt, Phys. Rev. **107**, 995 (1957); R.K. Nesbet and R.E. Watson, Ann.Phys. (N.Y.) **9**, 260 (1960); L.M. Sachs, Phys. Rev. **117**, 1504(1960); R.E. Watson and A. J. Freeman, Phys. Rev. **210**, 1125, 1134 (1960).
15. See e.g. P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955); H. Shull and P.-O. Löwdin: J. Chem. Phys. **25**, 1035 (1956).
16. C.C.J. Roothaan, Rev.Mod. Phys. **32**, 179 (1960).
17. P.-O. Löwdin, J.Appl.Phys.Suppl. **33**, 251 (1962).
18. P.-O. Löwdin, Rev. Mod. Phys. **35**, 496 (1963); Adv. Chem. Phys. **14**, 283 (1960).
19. D. J. Thouless: *The Quantum Mechanics of Many-Body Systems* (Academic Press, New York, 1961); W. Adams: Phys. Rev. **127**, 1650 (1962); J. Cizek and J. Paldus: J. Chem Phys. **47**, 3976 (1967); J. Paldus and J. Cizek: Prog. Theor. Phys. **42**, 769 (1969); J. Chem. Phys. **52**, 2919 (1970); J. Cizek and J. Paldus: J. Chem. Phys. **53**, 821 (1970); J. Paldus and J. Cizek: J. Polym. Sci. Part C **29**, 199 (1970); J. Paldus and J. Cizek: Phys. Rev. **A 2**, 2268 (1970); J. Paldus and J. Cizek: J. Chem. Phys. **54** (1971); J. Paldus, J. Cizek, and B. A. Keating: Phys. Rev. **A 8**, 640 (1973); A. Laforgue, J. Cizek, and J. Paldus: J. Chem. Phys. **59**, 2560 (1973); W. G. Laidlaw: Int. J. Quantum Chem. **7**, 87 (1973); J. Paldus and A. Veillard: Chem. Phys. Lett. **50** (1977); J. Paldus, J. Cizek, and A. Laforgue: Int. J. Quantum Chem. **13**, 41 (1978); J. Paldus and A. Veillard: Mol. Phys. **35**, 445 (1978); M. Benard and J. Paldus: J. Chem. Phys. **72**, 6546 (1980); H.

Fukutome: Prog. Theor. Phys. **40**, 998 (1968); **49**, 22 (1973); **50**, 1433 (1973); **52**, 115 (1974); **52**, 1766 (1974); **53**, 1320 (1975); M. Ozaki and H. Fukutome: Prog. Theor. Phys. **60**, 1322 (1978); M. Ozaki: Prog. Theor. Phys. **62**, 1183 (1979); M. Ozaki: Prog. Theor. Phys. **63**, 84 (1980); H. Fukutome: Int. J. Quantum Chem. **20**, 955 (1981).

20. P.-O. Löwdin, J.-L. Calais, and J. Calazans, Int. J. Quantum Chem. **20**, 1201 (1981); P. O. Löwdin: Proc. Ind. Acad. Sci. (Chem. Sci.) **96**, 121 (1986).
21. E. Lieb and B. Simon: Comm. Math. Phys. **53**, 185 (1977).
22. E. Schröder, Math. Ann. **2**, 317 (1870). See also P.-O. Löwdin, J.Mol.Spec. **10**, 12 (1963), particularly Appendix A.
23. A. W. Overhauser, Phys. Rev. Letters **3**, 414 (1959); *ibid.* **4**, 415. 462 (1960); W. Kohn and S. J. Nettel, Phys. Rev. Letters **5**, 8 (1960); K. Sawada and N. Fukuda, Prog. Theor. Phys. (Kyoto) **25**, 653 (1961); E. M. Henley andTh. W. Ruijgrok, Ann. Phys. (N.Y.) **12**, 409 (1961); E. M. Henley and L. Wilets, ibid. **14**, 120 (1961); A. W. Overhauser, Phys. Rev. **128**, 1437 (1962).
24. For a survey of some applications to solid-state theory, see e.g. J.-L. Calais, Adv. Quantum Chem. **17**, 225 (Academic Press, 1985).
25. T. A. Kaplan and W. H. Kleiner, Bull.Am. Phys. Soc. **11**, 234 (1966).
26. H. Fukutome: Prog. Theor. Phys. **40**, 998 (1968); **45**, 1382 (1971); **47**, 1156 (1972); **49**, 22 (1973); **50**, 1433 (1973); **52**, 115 (1974); **52**, 1766 (1974); **53**, 1320 (1975); M. Ozaki and H. Fukutome: Prog. Theor. Phys. **60**, 1322 (1978); M. Ozaki: Prog. Theor. Phys. **62**, 1183 (1979); M. Ozaki: Prog. Theor. Phys. **63**, 84 (1980); H. Fukutome: Int. J. Quantum Chem. **20**, 955 (1981); 27. P. A. M. Dirac, *The Principles of Quantum Mechanics* , 4th ed. (Clarendon, Oxford 1978).
28. P.-O. Löwdin, Ark. Mat. Astr. Fys. **30**, 1 (1948); J. Chem. Phys. **18**, 365 (1950); Adv. Phys. **5**, 1 (1956); Adv. Quantum Chem. **5**, 185 (Academic Press, New York, 1970).
29. See e.g. P.-O. Löwdin, Phys. Rev. **139**, A357 (1965): J. Chem. Phys. **43**, Suppl., S175 (1965).
30. D. R. Hartree: "*The Calculation of Atomic Structures*" (John Wiley and Sons, New York, 1957).
31. See e.g. P.O. Löwdin: in "NAS-ONR Report Shelter Island Conference 1951," 187 (1951); Quart. Appl. Math. **10**, 97 (1952).
32. R. S. Mulliken: Phys. Rev. **32**, 186 (1928).F. Hund: Z. Physik **51**, 759 (1928); see also J. E. Lennard-Jones, Trans. Faraday Soc. **25**, 668 (1929). as well as F. Bloch: Z. Physik **52**, 335 (1929); **57**, 545 (1929).

33. R. S. Mulliken: J. Chem. Phys. **46**, 497, 675 (1949).
34. C. C. J. Roothaan: Revs. Mod. Phys. **23**, 69 (1951). Compare also C. A. Coulson: Proc. Cambridge Phil. Soc. **34**, 204 (1938), as well as J. Lennard-Jones: Proc. Roy. Soc. London **A202**, 155 (1950).
35. S. F. Boys: Proc. Roy. Soc. (London), **A200**, 542 (1950); Svensk Kem. Tidskr. **67**, 367 (1955); Proc. Roy. Soc. (London), A258, 402 (1960); Revs. Mod. Phys. **32**, 296 (1960); see also Boys, Cook, Reeves and Shavitt, Nature **178**, 1207 (1956); S. F. Boys and G. B. Cook: Revs. Mod. Phys. **32**, 285 (1960).
36. C. A. Coulson and H. C. Longuet-Higgens: Proc. Roy. Soc. (London) **A191**, 39; 192, 16 (1947); **193**, 447, 456 (1948); **195**, 188 (1948).
37. P. O. Löwdin: Phys. Rev. **97**, 1490 (1955); R. McWeeny: Proc. Roy. Soc. London **A223**, 63, 306 (1954).
38. J. Musher, Chem. Phys. Lett. **7**, 397 (1970).
39. S. Lunell, Chem. Phys. Lett. **13**, 93 (1972), particularly p.96; G. Howat and S. Lunell, in *Quantum Science, Methods and Structure*, (Eds. J.-L. Calais et. al., Plenum Press, New York 1976), particularly p. 495.
40. H. Fukutome, J. Mol. Structure (Theochem) **188**, 337(1989).
41. W.J. Hehre, W.A. Lathan, R. Ditchfield, M.D. Newton, and J.A. Pople, GAUSSIAN-70, QCPE Program No. 236 (1970).
42. J.C. Slater, Phys. Rev. **35**, 509 (1930).
43. I. Mayer, Acta Phys. Hung. **54**, 249 (1983).
44. F.J. Lovas, J. Phys. Chem. Ref. Data **11**, 251 (1982).
45. I. Mayer, Chem.Phys. Letters **89**, 390 (1982).
46. P. O. Löwdin: Phys. Rev. **97**, 1509 (1955); Ann. Acad. Reg. Sci. Upsaliensis **2**, 127 (1958); Rev. Mod. Phys. **32**, 328 (1960); *Quantum Theory of Atoms, Molecules, and Solid-State; a Tribute to John C. Slater*, 601 (Ed. P.-O. Löwdin, Academic Press, New York 1966).
47. P.-O. Löwdin, Proc. Symp. Mol. Physics at Nikko, Japan, 1953, 13 (Maruzen, 1954); Proc **10**th Solvay Conf., *Les Electrons dans le Metaux*, 71 (Bruxelles 1955); Svensk Kemisk Tidskr. **67**, 365 (1955); T. Itoh and H. Yoshizumi: J. Phys. Soc. Japan **10**, 201 (1955); J. Chem. Phys. **23**, 412 (1955); Busseiron Kenkyu **83**, 13 (1955); P.-O. Löwdin, Texas J. Science **8**, 163 (1956); Ed. du Centre Nat. Rech. Sci. **82** (1958); Revs. Modern Phys. **32**, 328 (1960); R. Lefebvre, H. H. Dearman and H. M. McConnell: J. Chem. Phys. **32**, 176 (1960); P. O. Löwdin, R. Pauncz and J. de Heer: J. Chem. Phys. **36**, 2247, 2257, (1962); R. Pauncz: J. Chem. Phys. **37**, 2739 (1962); J. de Heer: Rev. Mod. Phys. **35**, 631 (1963); R. Pauncz: in *Molecular Orbitals in Chemistry, Physics, and Biology*, 433 (ed. P. O. Löwdin, Academic Press, New York, 1964) ; Tetrahedron **19**, Suppl.

2, 43 (1963); J. Chem. Phys. **43**, S69 (1965); O. Goscinski and J. L. Calais: Arkiv Fysik **29**, 135 (1965); J. de Heer and R. Pauncz: J. Chem. Phys. **39**, 2314 (1963); R. Pauncz: *Alternant Molecular Orbital Method*, (W. B. Saunders, Philadelphia, 1967); J. L. Calais: Arkiv Fysik; **28**, 479, 511, 539 (1965); **29**, 255 (1965); J. L. Calais: Int. J. Quantum Chem. **13**, 661 (1967). For more complete references, see I. Mayer: Adv. Quantum Chem. **12**, 189 (Academic Press, New York, 1980).

48. S. Lunell, Phys.Rev. **A 1**, 360 (1970); Chem. Phys. Lett. **13**, 93(1972); S. Lunell and G. Howat, in *Quantum Science, Methods and Structure*, 491, (Eds. J.-L. Calais et.al., Plenum Press, New York 1976)
49. N. H. F. Beebe and S. Lunell, J.Phys. B. **8**, 2320 (1975).

MANY CENTER AO INTEGRAL EVALUATION USING CARTESIAN EXPONENTIAL TYPE ORBITALS (CETO's)[1]

Ramon Carbó and Emili Besalú
Institute of Computational Chemistry
Faculty of Sciences
University of Girona
17071 Girona (Spain)

Keywords: Many Center AO Integrals. Molecular Basis Sets. ETO. STO. CETO.

[1] A contribution of the Grup de Química Quàntica del Institut d'Estudis Catalans.

List of Contents

1. Introduction: Brief Historical Background,General Intentions and Method

A few initial words seem necessary before the unfolding of this paper. In this section a brief historical sketch of the ETO integral problem is given first. After this initial focusing on the subject three main points will be presented. The general intentions, which have started the generation of this research, a schematic methodological survey on how the work is developed and some description of the ways the present account is written, taking the form of a final section remarks are presented here.

1.1. Short History

It would be a preposterous attempt to sum up in a simple paper like the present one, the immense effort made around the area of AO

integral computation over exponential type orbitals, ETO, since the early days of Quantum Chemistry. To describe in full detail the historical development in this field would occupy the time and space of a long paper or perhaps of a book. Thus, the history on this subject presented below will omit ineluctably many valuable contributions, which appeared to the authors not so relevant for the understanding of the present work.

As a foreword it must be said, perhaps constructing a too late homage to the brilliant contribution of professor Boys to Quantum Chemistry, that the first description of cartesian exponential type orbitals (CETO) was made thirty years ago by Boys and Cook [1]. One can probably think this fact as a consequence of the evolution of Boys's thought on the basis set problem and to the incipient ETO-GTO dilemma, which Boys has himself stated ten years earlier [2a].

Starting the schematical historic outline now, one can suppose ETO integral calculation starts with the pioneering work of Hylleraas [3], Kemble and Zener [4], Bartlett [5], Rosen [6], Hirschfelder [7], Coulson [8] and Löwdin [9], solidifying as a quantum chemical discipline with the publication of overlap formulae and tables by Mulliken et al. [10].

After this initial stage, the development of many ideas in the fifties by Kopineck [11], Hirschfelder and Linnett [12], Lundquist and Löwdin [13], Ruedenberg [14], Barnett and Coulson [15], Löwdin [16], Roothaan [17] and Hamilton [18], has opened the way to a full discussion on this topic in the sixties, where researchers continued to work on integrals over ETO functions. Among many, one can note in

this period the papers of: Shavitt and Karplus [19], Lofthus [20], Corbató and Switendick [21], Harris [22], Kotani et al. [23], Pitzer and Lipscomb [24], Fraga [25], Ruedenberg et al. [26] and Silverstone [27].

The interest on this subject has been extended into the seventies with the studies of Löwdin [28], Bosanac and Randić [29], Whitten [30], Sharma [31], Weatherford and Jain [32], who have contributed to the general development and understanding of the practical integral structure over ETO functions.

Recently, two points must be mentioned: a broad contribution of many authors which has been published by Weatherford and Jones as editors [33], and the clever treatment of Fernández-Rico et al. [34] where many references on the subject can be found. Without neglecting other meritorious pieces of research, in modern times the continued work of Guseinov [35], Jones [36] and Steinborn [37] must be taken into account and quoted.

On the other hand and in terms of an introductory excuse, let us describe some attempts to develop a naïve computational structure, based on an old idea of Ruedenberg [38], which have been made by one of us [39] and then, applied to approximate Electrostatic Molecular Potentials [40a] as well as to elucidate the origin of Mulliken's Magic Formula [40a] and to discuss in a general context various quantum chemical problems [40c]. This previous experience has promoted the present incursion on molecular integral computation, which has a recently published prologue [41].

1.2. General Intentions

A procedure will be presented here, inspired by Ruedenberg's [38] ingenuity and Saunders [42] foresight, whose roots can be found in the mathematical development of the specified preceding papers.

Despite the interest to obtain AO integral algorithms over cartesian exponential orbitals or STO functions [43] in a computational universe dominated by GTO basis sets [2], this research was started as a piece of a larger project related to Quantum Molecular Similarity [44], with the concurrent aim to have the chance to study big sized molecules in a SCF framework, say, without the need to manipulate a huge number of AO functions.

In the present paper the cartesian exponential orbital functions, CETO, structure will be defined first as well as their STO connection. In a second place will be deeply analyzed the subject on how to represent a product of a couple of these functions, centered at different points in space. Various practical examples on integral computation will follow, using some particular three and four center integrals, but also performing a general review over the most necessary related formulae. Detailed information on the structure of the algorithms used in all integral evaluation steps will be also given.

The structure of the formulation throughout the following paper has been constructed, as far as the authors capability permitted, in such a way as to allow an easy translation to the usual high level programming languages. In order to accomplish this objective, which is from the practical point of view capital for the future use of the

material described here, some new symbolic concepts connecting programming structures with mathematical formalism have been specially described in sections 2 and 3, and used afterwards in the subsequent discussions and integral formulae. More details on this subject can be found in reference [45b]

1.3. Method

The approach taken here, is based essentially on three main lines, once CETO functions had been chosen as the appropriate basis set:

1. Trivial integration over spherical and prolate spherical coordinate systems.
2. Use of overlap integrals and electrostatic potentials, essentially nuclear attraction integrals, when dealing with two electron repulsion integrals.
3. Development of a separate center density approach to solve the three and four center integral problem.

However, the authors do not claim that these three main strategic lines in company of CETO functions constitute the unique way nor the best path to solve the molecular integral problem directed to find plausible substitutes of GTO functions. Other integration methods to deal with the present discussion can be used and analyzed, for instance: Fourier, Laplace or Gauss transform methodology or any other possible choices and techniques available in the modern mathematical panoply.

Even other ETO forms can be studied as better basis sets, perhaps. To examine all these alternatives is not the purpose of the present paper, but one can consider the main goal of this work based on the task of digging out as much as possible information from CETO functions, regarded as STO building blocks.

1.4. Final Remarks

It must be stressed, throughout the development of the following discussion, the fact, besides the proper technical and mathematical aspects, that there has been a conscious intention, coming from the authors part, to provide details which perhaps can appear superfluous to the connoisseur's unaided eye.

Therefore, the authors must confess that, using sometimes simple terms and elementary arguments, they pretend to describe not only a piece of research circling again around the ETO integral problem, but they also want to produce a selfcontained text, which can serve to introduce students, as easily as possible while following a very general way, into the somehow difficult and arid territory of molecular integral calculation, and all this far away from the well known, literature dominating, GTO functions integral field.

An intended useful work, covering all the aspects of the problem, will be presented in the next sections.

2. Exponential Type Orbitals

A short introduction on Exponential Type Orbitals (ETO) functions is given here, in order to prepare the integral evaluation over cartesian ETO's.

Also, the connection of various ETO forms will be discussed, in particular details over the structure of Slater Type Orbitals (STO) [43a,46] with respect to CETO functions will be given.

All of this turning around the presentation and justification of the main subject of this paper, cartesian ETO functions: CETO.

In fact, the interest of CETO functions lies in their own potential to appear as an easy, obvious, alternative to Gaussian Type Orbitals (GTO) [2], the overwhelming functional structure, heading the vast majority of quantum chemical calculations of the present times.

2.1. CETO

Definition 1: CETO functions.

From now on let us define a real normalized Cartesian Exponential Type Orbital (CETO) centered at the point **A** as:

$$\chi(\mathbf{A},\mathbf{a},n,\alpha)= N_n(\mathbf{a},\alpha)\, H_a(\mathbf{A},\mathbf{a})\, P_n(\mathbf{A},\alpha). \qquad (2.1)$$

where the following conditions hold:

1) $\mathbf{a}=(a_1,a_2,a_3)$, a_i integer ≥ 0 $\forall i$;
2) n integer ≥ 0;
3) α real >0;
4) $r_A=|\mathbf{r}-\mathbf{A}|$, $\mathbf{r}=(x_1,x_2,x_3)$, $\mathbf{A}=(A_1,A_2,A_3)$.

The terms appearing in equation (2.1) are the following:

a) The **normalization factor**:

$$N_n(\mathbf{a},\alpha) = \alpha(2\alpha)^{a+n+\frac{1}{2}} [(2a+1)!! \ (\pi \ p \ (2(a+n+1))!)^{-1}]^{\frac{1}{2}}, \qquad (2.2)$$

where $a=\Sigma_i a_i$ and $p=\prod_i (2a_i-1)!!$.

b) The **angular part**:

$$H_{\mathbf{a}}(\mathbf{A},\mathbf{a}) = \prod_i (x_i-A_i)^{a_i}, \qquad (2.3)$$

where a is defined as in equation (2.2) and can be called the **order of the angular factor**.

c) The **radial part**:

$$P_n(\mathbf{A},\alpha) = r_A^n \exp(-\alpha r_A), \qquad (2.4)$$

where n can be taken as the **order of the radial factor**, and α is an exponential scale factor.■

In this manner, one can see how CETO functions are a particular case of a more general radial part:

$$P_n(\mathbf{A},\alpha,k) = r_A^n \exp(-\alpha r_A^k), \tag{2.5}$$

so if k=1 one describes a CETO and a GTO corresponds to k=2.

In some cases condition 2) above can be relaxed allowing negative values, and condition 3) may be extended with $\alpha=0$. The relaxation-extension of these two conditions permits some operators to be expressed as CETO functions.

This CETO form definition has the advantage, among others, to permit the construction of real cartesian STO's easily, see for example [42,46] as a particular introduction to STO cartesian functions. A more detailed analysis of such a possibility follows.

2.2. STO

Slater type orbital (STO) functions will be presented here. They are well known functions since their description made by Slater around sixty years ago [43a] and heavily studied and used as atomic basis sets [43b,43c]. In this section the connection of STO with CETO functions is analyzed.

STO functions are essentially complex functions represented in a classical manner within a spherical coordinate system. Despite these characteristics it is not difficult to relate them with real functions

described in a cartesian coordinate system as CETO functions are intentionally constructed.

2.2.1. Complex STO over Spherical Coordinates

Definition 2: STO functions

Let us consider, for example, the convenient Wahl et al. definition [47] of normalized Slater type orbitals (STO) constructed over spherical coordinates and referred to the origin:

$$S(n,\ell,m,\alpha) = N_n(\alpha)\, R_n(r,\alpha)\, Y_{\ell m}(\theta,\varphi), \tag{2.6}$$

where the following conditions hold:

1) n integer >0;
2) $0 \leq \ell \leq n-1$;
3) $-\ell \leq m \leq +\ell$;
4) α real >0.■

The three terms in equation (2.6) can be built up by means of the same parts as CETO functions have been constructed before:

a) The **normalization factor**:

$$N_n(\alpha) = (2\alpha)^{n+\frac{1}{2}}\, ((2n)!)^{-\frac{1}{2}}. \tag{2.7}$$

b) The **angular part** defined as a spherical harmonic function [47,48,49]:

$$Y_{\ell m}(\theta,\varphi) = P_{\ell m}(\cos\theta)\, \Phi_m(\varphi), \tag{2.8}$$

where

$$\Phi_m(\varphi) = (2\pi)^{-1/2}\, e^{im\varphi}, \tag{2.9}$$

and the **associated Legendre polynomials** [50] can be written as:

$$P_{\ell m}(x) = M(\ell,m)\, Q_{\ell m}(x), \tag{2.10}$$

with the normalization factor:

$$M(\ell,m) = \left(((2\ell+1)/2)\, ((\ell-m)!/(\ell+m)!) \right)^{1/2}. \tag{2.11}$$

The structure of an **associated Legendre polynomial** is:

$$Q_{\ell m}(x) = (-1)^m\, (1-x^2)^{m/2} \sum_{\mu=0}^{M} w_\mu^{\ell m}\, x^{\ell-m-2\mu}, \tag{2.12}$$

with the additional definitions of the upper summation value defined as the greatest integer in the quotient $(\ell-m)/2$: $M=[\![(\ell-m)/2]\!]$, and the polynomial coefficients made by:

$$w_\mu^{\ell m} = (-1)^\mu\, (2\ell-2\mu)!/\, \{2^\ell\, (\ell-m-2\mu)!\, (\ell-\mu)!\, \mu!\}. \tag{2.13}$$

The STO structure described above apples to non-negative m values. When considering negative values of m it must be taken into account that the following relationship holds within the spherical harmonic function set:

$$Y_{\ell,-m}(\theta,\varphi) = (-1)^{|m|}\ Y^{*}_{\ell,|m|}(\theta,\varphi). \tag{2.14}$$

c) The **radial part**, written as:

$$R_n(r,\alpha) = P_{n-1}(0,\alpha) = r^{n-1}\ e^{-\alpha r}. \tag{2.15}$$

The same remarks as in CETO functions description can be also considered here.

After this classical STO formulation, here must be noted the coincidence of s-type STO and some particular forms of CETO functions. These s symmetry functions can be represented by $S(n,0,0,\alpha)$ in the STO function framework and by the origin centered equivalent CETO: $\chi(0,0,n-1,\alpha)$ forms, respectively.

2.2.2. STO's over Cartesian Coordinates

a) Remembering the previous definitions as well as the relationships between cartesian and spherical coordinates, which can be written as:

$$\begin{aligned} z\ r^{-1} &= \cos\theta, \\ (x+iy)\ r^{-1}\ e^{-i\varphi} &= \sin\theta, \end{aligned} \tag{2.16}$$

after some straightforward simplifications one can write an all purpose general cartesian unnormalized STO (GC-STO) function as:

$$S_G(n,\ell,m,\alpha) = P_{n-(\ell+1)}(\mathbf{A},\alpha)\; C_G(\ell,|m|,\mathbf{r}-\mathbf{A}), \qquad (2.17)$$

where:

$$C_G(\ell,|m|,\mathbf{A}) = \sum_{\mu=0}^{M} \{ (w_\mu^{\ell|m|}\; r_A^{2\mu}) \bullet \sum_{\tau=0}^{M} \{ (-1)^\tau\, Z(|m|,G(\tau))\; H_{\ell-2\mu}(\mathbf{A},\mathbf{a}(\ell,|m|,\mu,G(\tau)) \} \}, \qquad (2.18)$$

using the fact that: $r_A^2=x_A^2+y_A^2+z_A^2$, defining the vector: $\mathbf{a}(\ell,|m|,\mu,G(\tau))=(|m|-G(\tau),\; G(\tau),\; \ell-(|m|+2\mu))$, and the upper summation limits taken as the largest integers in the bracketed quotients: $M=[\![(\ell-|m|)/2]\!]$ and $N=[\![|m|/2]\!]$.

The GC-STO angular part coefficients $Z(|m|,G(\tau))$ of equation (2.18) are defined by the expression:

$$Z(|m|,G(\tau)) = \binom{|m|}{G(\tau)}\, \delta(G(\tau)\leq |m|), \qquad (2.19)$$

where $G(\tau)$ is a simple function of τ and use is made of a logical Kronecker delta, defined as follows:

Definition 3: Logical Kronecker delta.

A logical Kronecker delta $\delta(\{L\})$ is a symbolic factor which is unity when the logical expression {L} is true and zero otherwise.■

Frequent use will be made of this Definition 3 in the text that follows. An initial and crude form of this symbol was given some years ago by one of us within the development of a MCSCF context [45a], and it has been recently studied in deep in our Laboratory [45b].

b) After this description of the GC-STO functions, a complex cartesian STO form can be easily build up:

$$S(n,\ell,m,\alpha) = N(n,\ell,m,\alpha)\, P_{n-(\ell+1)}(A,\alpha) \bullet \{ C_R(\ell,|m|,A) + (-1)^{\delta(m<0)}\, i\, C_I(\ell,|m|,A) \}, \qquad (2.20)$$

with a normalization factor, collecting a sign constant and all the various norms appearing in the STO Definition 2 above, which is written as:

$$N(n,\ell,m,\alpha) = (-1)^{\delta(m>0 \wedge m\neq 2)}\, N_n(\alpha)\, M(\ell,|m|)\, (2\pi)^{-1/2}. \qquad (2.21)$$

The function $G(\tau)$ appearing in the C coefficients, as can be found in equation (2.19), is defined to use into the real part as $R(\tau)=2\tau$ and for the same purpose is taken into the imaginary one as $I(\tau)=2\tau+1$.

c) In any case the real and imaginary functions: C_R and C_I, conveniently multiplied by a radial function can be taken as real unnormalized STO functions in cartesian form. The interesting pattern here is the appearance of the structure of cartesian STO functions, complex or real, expressible as linear combinations of CETO's. This property can be also used in similar contexts where spherical harmonic functions are present. Take as an example the atomic spinor functions, see for instance references [48]. It will be now easy to design such a relativistic one electron orbitals as formed by four dimensional vectors whose components are CETO linear combinations. From here one can imagine, perhaps, an interesting starting point towards constructing the road in the direction of high quality relativistic SCF or CI molecular calculations.

2.2.3. Corollary

There are some benefits in using STO instead of GTO functions, the problem consists on how to solve the many center integral computation bottleneck. A possible way will be discussed now, using CETO's.

As it is shown above, CETO functions can be considered the building blocks of cartesian STO's. In this sense, solving the many center integral evaluation problem over CETO functions will be the same as to move another step into the path to solve the integral STO problem.

The point to consider at this moment is: being STO's linear combinations of CETO functions, one can propose the building blocks themselves as variational probe functions instead of the more complicated cartesian STO expressions. CETO functions become in this manner, as GTO functions are, firm candidates to stardom instead of simple second rank guest actors.

2.3. Other Possible ETO Forms

Because of some important relationship with CETO functions on the one hand which will be used later on, and due to their interesting structural properties on the other hand, two alternative ETO forms will now be discussed.

2.3.1. SETO

Within the framework of spherical coordinates $\{r_A,\theta_A,\varphi_A\}$ one can define an unnormalized **Spherical Exponential Type Orbital** (SETO) function as:

$$S(\mathbf{A},\mathbf{a},n,\alpha) = J_a(\mathbf{A},\mathbf{a})\ P_n(\mathbf{A},\alpha), \tag{2.22}$$

where $a=\Sigma_i a_i$ is the order of the **angular part** defined in terms of spherical coordinates as:

$$J_a(\mathbf{A},\mathbf{a}) = (\sin\theta)^{a_1}\ (\cos\theta)^{a_2}\ (\sin\varphi)^{a_3}\ (\cos\varphi)^{a_4}, \tag{2.23}$$

This suggests, as sine and cosine functions are trivially related, that perhaps one can define some sine and cosine SETO forms with angular parts written as follows:

$$L_s(\mathbf{A},a,b) = (\sin\theta_A)^a\ (\sin\varphi_A)^b, \tag{2.24}$$

or

$$K_s(\mathbf{A},a,b) = (\cos\theta_A)^a\ (\cos\varphi_A)^b, \tag{2.25}$$

with s=a+b. Thus, one can rewrite the SETO structure, using the CETO radial part defined in equation (2.4), as:

$$S(\mathbf{A},a,b,c,\alpha) = L_s(\mathbf{A},a,b)\ P_c(\mathbf{A},\alpha), \tag{2.26}$$

or

$$C(\mathbf{A},a,b,c,\alpha) = K_s(\mathbf{A},a,b)\ P_c(\mathbf{A},\alpha), \tag{2.27}$$

which will act as elementary ETO building blocks in spherical coordinates. SETO functions may be interesting to perform atomic

calculations and obviously are the most suitable ETO form to compute one center integrals.

The importance of SETO functions may be in some aspects greater than CETO forms, because CETO's can be written in terms of SETO functions, thus the last ones constitute a still more general AO building block. As an example, one can easily see that a CETO function as $\chi(\mathbf{A},\mathbf{a},n,\alpha)$ corresponds to a SETO which has the following structure:

$$\chi(\mathbf{A},\mathbf{a},n,\alpha) = S(\mathbf{A},\mathbf{b},n+a,\alpha) = J_t(\mathbf{A},\mathbf{b})\, P_{n+a}(\mathbf{A},\alpha), \tag{2.28}$$

with $a=\Sigma_i a_i$, $t=2a-a_3$ and being the four dimensional vector **b** defined in terms of their components:

$$\mathbf{b} = (a_1+a_2,a_3,a_2,a_1). \tag{2.29}$$

An also interesting SETO function feature corresponds to the separability of these function structure in terms of the three spherical coordinates as STO functions are. Equation (2.22) may be written using still a partition of the angular part (2.23) into two functions:

$$Z_z(\mathbf{A},\mathbf{z}) = (\sin\theta_A)^{z_1} (\cos\theta_A)^{z_2}, \tag{2.30}$$

and

$$F_f(\mathbf{A},\mathbf{f}) = (\sin\varphi_A)^{f_1} (\cos\varphi_A)^{f_2}, \tag{2.31}$$

where $z=z_1+z_2$ and $f=f_1+f_2$, thus:

$$J_a(\mathbf{A},\mathbf{a}) = Z_z(\mathbf{A},\mathbf{z})\, F_f(\mathbf{A},\mathbf{f}), \tag{2.32}$$

with $a=z+f$ and $\mathbf{a}=\mathbf{z}\oplus\mathbf{f}$

2.3.2. LETO

Another kind of ETO functions, which also can be formed having a cartesian term as angular component, are the functions which may be called **Laplace Exponential Type Orbitals**: LETO. An unnormalized LETO can be defined as follows:

$$L(\mathbf{A},\mathbf{a},\alpha) = H_a(\mathbf{A},\mathbf{a})\, Q(\mathbf{A},\alpha), \tag{2.33}$$

where $H_a(\mathbf{A},\mathbf{a})$ has the same definition as in the equation (2.3) and $Q(\mathbf{A},\alpha)$ is the term that confers on the LETO functions some exciting peculiarities. Its definition may be written as:

$$Q(\mathbf{A},\alpha) = \exp\{-\Sigma_i \alpha_i \,|\, (x_i - A_i) \,|\,\}, \tag{2.34}$$

and shows the fact that LETO functions have one of the most characteristic properties also encountered in the GTO functional structure: the separability into three unidimensional terms associated to each one of the cartesian coordinates, that is:

$$L(\mathbf{A},\mathbf{a},\alpha) = \prod_i \Lambda(A_i,a_i,\alpha_i), \tag{2.35}$$

where

$$\Lambda(A_i,a_i,\alpha_i) = (x_i - A_i)^{a_i}\, q(A_i,\alpha_i), \tag{2.36}$$

with each monodimensional exponential term written as a **Laplace distribution**:

$$q(A_i,\alpha_i) = \exp\{-\alpha_i \,|\, (x_i - A_i) \,|\,\}. \tag{2.37}$$

2.3.3. Remarks

Although SETO and LETO functions offer characteristic peculiarities, which are lacking in **cartesian** STO and CETO functions, namely: complete separability in spherical and cartesian coordinates, respectively, their application into integral formulae calculation will demand a still larger paper than the present one. In this manner the authors consider to keep these challenging ETO forms unstudied for the moment is the best policy and propose accordingly to discuss here CETO integrals only, leaving aside, but prepared for a possible future exposition, SETO and LETO integral forms.

From now on only CETO functions will be considered.

3. Practical Details on CETO Products

This section will discuss the practical structure of the CETO functions being used, when included in a molecular basis set, to compute AO integrals for atomic and molecular calculations.

As it is well known, basis functions when used in a LCAO framework appear paired at the moment of computing integrals. The structure of the two centered pairs of CETO functions and their transformation behavior opens the way to describe simpler CETO forms, which may be employed, in turn, as integral building blocks.

The structure of these CETO products will allow the definition, at the end of this section, of canonic monoelectronic integral forms,

introducing the various kinds of auxiliary integrals, which will be used as basic computational entities later on in sections 6 and 7.

3.1. Orientation of CETO's over the Line Joining a Pair of Centers

a) A product of two CETO functions in order to be useful from the integral computation perspective must be **well-oriented**, in such a sense that the x_3-axis (z-axis), say, coincides with the line joining the two involved centers.

As also happens when dealing with STO functions it will be necessary to perform some CETO rotations bringing the molecular reference frame into the bicentric axis system.

Due to the general form of CETO functions as defined in equation (2.1), rotations need only to be done over cartesian coordinates contained in the CETO angular part. The variables appearing in the vector $\mathbf{r}$-$\mathbf{A}$ of the angular part will be rotated, but not those in the norm $r_A = |\mathbf{r}\text{-}\mathbf{A}|$ or any other function of r_A or r_B, as a consequence of the spherical symmetry attached to the vector radius.

b) Here it is analyzed in the first place how to perform an effective rotation over the coordinates $\mathbf{x}^o=(x_1^o,x_2^o,x_3^o)$ using the eulerian angles $\omega=(\varphi,\theta,\delta)$ and the rotation matrix $\mathbf{R}(\omega)$ [49]. The rotation $\mathbf{R}(\omega)$ is chosen in order to express the coordinates $\mathbf{x}^o$ in terms of the coordinate system $\mathbf{x}=(x_1,x_2,x_3)$ having the same origin as $\mathbf{x}^o$, **A** say, and which is well oriented with respect to another center, **B**.

c) The rotations needed in the present paper from now on will be used to well-orient two CETO functions, centered in different sites and described in a prolate spherical coordinates framework. For the detailed construction of such a coordinate system see Definition 7 in section 3.4.2.a below.

The involved rotation angles for this kind of manipulation will be in the present case constructed by the vector: $\omega_s=(\varphi,\theta,0)$. That is because of the cylindrical symmetry which the prolate spherical coordinates system possess with respect a rotation around the axis line joining both centers. This kind of rotation process form will be assumed throughout the remaining text.

The two Euler angles needed in order to construct a given rotation vector ω_s are computed using the following scheme:

Procedure 1: **Computation of Euler angles φ and θ.**

Given the vector $\mathbf{r}=(x_1,x_2,x_3)$ joining the two involved centers, then the Euler's angles (φ,θ) are computed using the simple algorithm:

$$\varphi=\tan^{-1}(x_2/x_1);\ \text{if } x_1<0 \text{ then } \varphi=\text{Mod}(\varphi+\pi,2\pi)$$
$$\theta=\cos^{-1}(x_3/|\mathbf{r}|).\blacksquare \qquad (3.1)$$

The orthogonal rotation matrix $\mathbf{R}(\omega_s)$ which will perform the rotation using an angle set ω_s is defined as:

$$\mathbf{R}(\omega_s)=\begin{bmatrix} \cos\varphi\cos\theta & -\sin\varphi & \cos\varphi\sin\theta \\ \sin\varphi\cos\theta & \cos\varphi & \sin\varphi\sin\theta \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}. \qquad (3.2)$$

Hence one can write:

$$\mathbf{x}^{o} = \mathbf{x}\, \mathbf{R}^{T}(\omega_{s}), \tag{3.3}$$

where $\mathbf{R}^{T}(\omega_s)$ means here the transpose of $\mathbf{R}(\omega_s)$.

d) Moreover, cartesian coordinate power products as these forming the CETO angular part: $H_a^o(\mathbf{A},\mathbf{a})$, are of special interest when dealing with this kind of AO functions. Let us show how a product expression is transformed under such a rotation.

Performing the matrix product in equation (3.3) and using the fact that:

$$(a+b+c)^k = \sum_{i=0}^{k} \binom{k}{i} c^{k-i} \sum_{j=0}^{i} \binom{i}{j} a^j b^{i-j}, \tag{3.4}$$

once obtained the triple coordinate product the following relationship can be found:

$$H_a^o(\mathbf{A},\mathbf{a}) = \Sigma(\mathbf{i}=0,\mathbf{a})\, \Sigma(\mathbf{j}=0,\mathbf{i})\, M(\mathbf{i},\mathbf{j},\mathbf{a})\, f(\mathbf{i},\mathbf{j},\mathbf{a},\varphi,\theta)\, H_a(\mathbf{A},\mathbf{g}), \tag{3.5}$$

where $a=\Sigma_i a_i=\Sigma_i g_i$ and some characteristic definitions, identities and remarks should be taken into account in order to interpret equation (3.5). The relationship with CETO functions angular part transformation will be given below in next f) paragraph. The most important of the symbolic conventions appearing in equation (3.5) can be defined as follows:

1. Definition 4: Nested Summation Symbols.

With respect to the symbol $\Sigma(\mathbf{k}=\mathbf{0},\mathbf{N})$, let us use from now on a general expression like $\Sigma(\mathbf{k}=\mathbf{r},\mathbf{s})$ as a sequence of sums and borrowing a programming terminology let us call it a **nested summation symbol**.

The meaning of such a convention corresponds to perform all the sums involved in the generation of all the possible forms of vector **k**, whose elements are defined by the limits $k_i=r_i,s_i;\forall i$. In this manner $\Sigma(\mathbf{k}=\mathbf{0},\mathbf{N})$ symbolizes all the nested sums leading to the construction of the complete **k** vector elements defined as: $k_i=0,N_i;\ \forall i$.

In order to use in practice this definition, it must also be decided which one of the nested sums will be considered the innermost. That is: the first index of vector **k** which is to be incremented in a first place. Here, in the present paper is followed the natural convention, this will be supposed to hold for the n-th sum symbol or same to say: for the **k** vector n-th index. Therefore, being the first vector index associated to the outermost nested sum, then it will be connected to the last index to suffer increment.

Still a more general nested summation symbol can be defined and written as: $\Sigma(\mathbf{k}=\mathbf{r},\mathbf{s},\mathbf{d})$, where the vector **d** contains all the arbitrary increments associated to the progress of each nested sum index.

In the cases where it is important to explicitly signal the **dimension of the nested sum**, that is the number of sums implied

in the symbol and thus, the dimension of all the vector indices involved, one can use a subindex on the nested summation symbol as: $\Sigma_n(\mathbf{k=r,s,d})$, meaning with this that n sums are involved in the nest.

When initial, final, increment and dimension values are implicit in the nested sum a simplified symbol such as $\Sigma(\mathbf{k})$ may be also used.■

As it can be deduced from Definition 4 above, the equivalence $\Sigma(\mathbf{k=r,s})\equiv\Sigma(\mathbf{k=r,s,1})$ holds, employing the symbol $\mathbf{1}=(1,1,...,1)$, because in all the formulae used in the remaining of this work sum index increments are unity.

It has also to be taken into account that a product of two nested summation symbols is another summation symbol. That is: $\Sigma_n(\mathbf{i})\Sigma_m(\mathbf{j})=\Sigma_{n+m}(\mathbf{k})$, where the vector $\mathbf{k}$ is constructed by the direct sum $\mathbf{k=i\oplus j}$. The rest of the nested sum constituent vectors are constructed in the same way.

Nested summation symbols applications have been deeply studied elsewhere [45b].

2. The coefficients M in equation (3.5) are defined by the binomial coefficient product:

$$\mathbf{M(i,j,a)} = (-1)^{a_1-i_1+i_3}\binom{a_1}{i_1}\binom{a_2}{i_2}\binom{a_3}{i_3}\binom{i_1}{j_1}\binom{i_2}{j_2}. \tag{3.6}$$

3. The function f in (3.5) is defined in turn as:

$$\mathbf{f(i,j,a},\varphi,\theta) = J_t(\mathbf{A,b}), \tag{3.7}$$

where the $J_t(\mathbf{A},\mathbf{b})$ function has been already described when defining SETO functions angular parts in equation (2.23) and $t=i_1+i_2+\Sigma_i a_i$ in the present case. The four dimensional vector **b** is built up in terms of vectors **a**,**i** and **j** elements as:

$$\mathbf{b} = \mathbf{b}(\mathbf{a},\mathbf{i},\mathbf{j}) = (i_1+i_2+i_3-(j_1+j_2),a_3+j_1+j_2-i_3,a_1+i_2-i_1,a_2+i_1-i_2). \qquad (3.8)$$

4. The three dimensional vector **g** components, appearing as the coordinate power exponents in (3.5) are constructed by the rules:

$$\mathbf{g} = \mathbf{g}(\mathbf{a},\mathbf{i},\mathbf{j}) = (j_1+j_2+i_3,a_1+a_2-(i_1+i_2),a_3+i_1+i_2-g_1). \qquad (3.9)$$

5. When using the present notation in this rotation case, the index j_3 must be considered constant and equal to zero. This particularity will supposed to be implicit in all the remaining text.

e) Some remarks on these previous five points are needed.

Nested summation symbolism is very convenient for practical purposes, because successive generation of **k** vector elements can be programmed in a general way under any high level language, using a unique **do** or **for** loop irrespective of the number of sums, or binomial expressions of type (3.4) entering the product (3.5).

For a discussion and an example on how to program an arbitrary number of nested loops using a single loop algorithm, allowing to realize an arbitrary number of nested sums, see references [45b,c].

A compact and short example in order to illustrate nested summations usefulness, programmed using a hybrid Fortran 90 high level language follows:

```
Program 1: Nested sum.

   comment < Programming example of a nested summation
             symbol: Σ(k=No,Nf,D). The innermost loop
             corresponds to the index i=n. Summa is a
             called procedure where the n nested loops
             converge and constructs a sum value s >

             Parameter n
             Data No(n),Nf(n),D(n)
             Dimension k(n)

   comment < Put initial values on index vector k >

             do i=1,n
                k(i)=No(i)
             end do

   comment < Start nested sum procedure >

             i=n
             do
                  if ( k(i)>Nf(i) ) then
                            k(i)=No(i)
                            i=i-1
                  else
                            call Summa(n,k,s)
                            i=n
                  end if
                  if (i≠0) then
                            k(i)=k(i)+D(i)
                  else
                            exit
                  end if
             end do

   comment < End of Program Nested Sum > ■
```

Nested Summation Symbols are ideal constructs for programming in a suitable hardware environment, for example: using transputer boards [51] or any computing device with various CPU running in parallel [52]. All operations attached to each vector **k** can be run in a separate CPU. Counting the times this can be done in

terms of the available independent processors is a trivial matter. Once nested sums are written in a similar form to the structure of Program 1, a counter can be defined which sends the corresponding index vector to the next idle CPU. The final sum can be accumulated in so many parts as processors are active in the system. Partial results are gathered from each CPU and summed up when the end of the nested sum is reached, see reference [45b] for more details.

Finally, it must be stressed that, as it will be shown later on in sections 6 and 7, where they are heavily used, nested summation symbols permit one to write integral formulae over CETO functions in a very compact manner.

f) The result encountered in equation (3.5) proves that any CETO placed at center **A** is a finite linear combination of a set of CETO functions also centered at **A**, which are well-oriented with respect another center though. The same occurs when STO functions are considered.

Function f appearing in equation (3.5) as defined in equation (3.7) is, in general, a constant due to the fact that the angles φ and θ are also constants. But in some integrals these angles can be considered variables, and the function f assumes an active integrand role as well. This feature will be encountered again, when dealing with some two-electron integrals in section 7.

From now on, let us simplify equation (3.5) when needed, writing the whole equation structure here in a symbolic manner:

$$H_a^o(\mathbf{A},\mathbf{a}) = \Sigma_i\, C_i^{AB}\, H_a(\mathbf{A},\mathbf{g}_i), \tag{3.10}$$

where the coefficients $M(\mathbf{i},\mathbf{j},\mathbf{a})$ and the mentioned function $f(\mathbf{i},\mathbf{j},\mathbf{a},\varphi,\theta)$, with φ and θ constants, are embedded in the coefficients C_i^{AB}. Here, the notation AB means that the axis system at center **A** is well oriented with respect to center **B**.

g) When two successive orientations are necessary, for example from the original axis system O to AB and from AB to AC say, one can describe the first rotation using equation (3.10). The transformed angular part can be afterwards expressed as:

$$H_a(\mathbf{A},\mathbf{g}_i) = \Sigma_j\, C_{ji}^{BC}\, H_a(\mathbf{A},\mathbf{a}_{ji}), \tag{3.11}$$

then substituting equation (3.11) into expression (3.10) it is obtained:

$$\begin{aligned} H_a^o(\mathbf{A},\mathbf{a}) &= \Sigma_i\, C_i^{AB}\, \Sigma_j\, C_{ji}^{BC} H_a(\mathbf{A},\mathbf{a}_{ji}) \\ &= \Sigma_j\, (\Sigma_i\, C_i^{AB}\, C_{ji}^{BC} H_a(\mathbf{A},\mathbf{a}_{ji})) \\ &= \Sigma_j\, C_j^{AC}\, H_a(\mathbf{A},\mathbf{a}_j). \end{aligned} \tag{3.12}$$

That is: the final result is a set of angular CETO parts well oriented with respect to the AC direction. This fact reflects the general properties of coordinate orthogonal transformations.

3.2. Well Oriented CETO Form

a) It is easy to consider how the general CETO structure, described in equation (2.1) of Definition 1 transforms and obtains a practical form. A Well-Oriented CETO (WO-CETO) function can be defined:

Definition 5: WO-CETO Functions.

Using the conventions:

1) $\mathbf{n}=(n_1,n_2,n_3)$;
2) $m \in \{0,1\}$;
3) $\alpha \geq 0$ and real;
4) $R \geq 0$ and real;
5) $s \in \{-1,0,+1\}$;
6) $z(s)=sR/2$;
7) $r(s)=(x_1^2 + x_2^2 + (x_3-z(s))^2)^{1/2}$;

an unnormalized WO-CETO function is written as:

$$W(z(s),\mathbf{n},m,\alpha) = x_1^{n_1}\, x_2^{n_2}\, (x_3-z(s))^{n_3}\, r(s)^m \exp\{-\alpha r(s)\}. \qquad (3.13)$$

One can represent unnormalized WO-CETO functions as a product of the following terms, when the convention $\mathbf{z(s)=(0,0,z(s))}$ and the definitions (2.3) and (2.4) apply:

a.1. An **angular part**:

$$H_n(\mathbf{z(s),n}) = x_1^{n_1}\, x_2^{n_2}\, (x_3\text{-}z(s))^{n_3}, \tag{3.14}$$

where $n=\Sigma_i n_i$.

a.2. A **radial part**:

$$P_m(\mathbf{z(s)},\alpha) = r(s)^m \exp\{-\alpha r(s)\}. \tag{3.15}$$

In this way $r_A=r(-1)$, so one can also use the symbol $W(A,\mathbf{n},m,\alpha)$ for $W(z(-1),\mathbf{n},m,\alpha)$; moreover one can call $r_B=r(+1)$ and consequently: $W(B,\mathbf{n},m,\alpha)=W(z(+1),\mathbf{n},m,\alpha)$.■

In the same manner as the n value constitutes an order of the WO-CETO angular part, the $m\in\{0,1\}$ value classifies WO-CETO, and also CETO functions, into two kinds: **odd** WO-CETO's if m=1 and **even** WO-CETO's if m=0. A similar classification may be applied over cartesian GTO functions, whose basis sets used in the literature are always defined as even.

Integrals over CETO functions can be described over WO-CETO's and transformed afterwards to the working molecular axis framework as needed. A similar procedure is used when dealing with STO functions, see for example [34e].

b) As equations (3.14) and (3.15) show, the WO-CETO structure (3.13) can be alternatively written, using a similar partition to the one discussed in equation (2.1) in Definition 1. A still simpler expression, taking the form of some sort of elementary CETO building block can be also described:

Definition 6: E-CETO Functions.

Taking into account the same conventions as in Definition 5, let us define an Elementary CETO (E-CETO) function as:

$$E(\mathbf{n},m,\alpha) = W(z(0),\mathbf{n},m,\alpha).\blacksquare \qquad (3.16)$$

E-CETO functions as WO-CETO ones can also be classified as odd or even, depending on the value of the integer m.

c) It must be noted here that the most convenient r_A, r_B exponent, from a computational point of view, in the above Definitions 1,5 and 6, is even as can be deduced by analogy from the GTO universal practice. The odd value has been also considered here in

order to produce a general formulation, but one must be aware that it may cause additional difficult patterns when developing some integral formulae. A similar problem is found when dealing with cartesian GTO's. This is simply due to the presence of a square root inherent to the vector radius definition.

d) With respect to this problem one can envisage a solution to overcome the CETO functions partition in two classes. In fact, it must be satisfactory enough possessing the ability to obtain a sufficiently good representation of an odd CETO radial part with respect a basis set of even ones, from the machine precision point of view.

In order to visualize this previous statement, consider that a simple equation as:

$$P_1 = r\, e^{-r} \approx e^{-ar} - e^{-br} = D_2, \qquad (3.17)$$

with two fitting parameters {a,b}, permits one to write the similarity measure between P_1 and D_2 functions as an integral:

$$Z = \int P_1\, D_2\, r^2\, dr, \qquad (3.18)$$

which after taking into account the normalization factors, has the simple expression:

$$Z = 12\, \{\, 3(4^{-1}(a^{-3}+b^{-3}) - 4(a+b)^{-3})\,\}^{-1/2}\ \{\, (1+a)^{-4} - (1+b)^{-4}\,\}. \qquad (3.19)$$

Over equation (3.19) it can be numerically tested that there is an infinite collection of optimal pairs of parameters, for example a=0.99792 and b=1.00208, for which: $Z \approx 1$, within the machine accuracy. That is because if $P_1=D_1$, then an infinite number of equations such as: $b^k-a^k=k$; $\forall k \geq 1$ hold.

Probably, an expansion like:

$$D_n = \sum_{i=1}^{n} c_i \, r^{2p_i} \exp(-a_i r), \tag{3.20}$$

where $\{c_i,p_i,a_i\}$ are variational parameters, can add an additional flexibility to the covering expression of P_1 by means of even CETO functions.

Another interesting possibility, which will not be discussed here, may consist into expanding the radial factors e^{-r} and re^{-r} as a linear combination of primitive gaussian functions [69]. Studies about this simpler approximation is under way in our Laboratory.

All the discussions performed in section 5 below may be taken as to have a general application, although only even CETO's are involved.

3.3. General CETO Products and Their Properties

a) One can find, when two CETO functions share the same center, their product is but a new CETO within the same center, or:

$$\Omega_{AA'} = \chi(\mathbf{A},\mathbf{a},n,\alpha)\ \chi(\mathbf{A},\mathbf{a}',n',\alpha') = \chi(\mathbf{A},\mathbf{a}+\mathbf{a}',n+n',\alpha+\alpha'). \qquad (3.21)$$

b) The product of two CETO's, placed respectively at centers **A** and **B**, separated a distance R apart, giving a charge distribution Ω_{AB} when rotated, can be written as:

$$\begin{aligned}\Omega_{AB} &= \chi(\mathbf{A},\mathbf{a},n,\alpha)\ \chi(\mathbf{B},\mathbf{b},m,\beta)\\ &= \Sigma(\mathbf{i}=\mathbf{0},\mathbf{a})\ \Sigma(\mathbf{j}=\mathbf{0},\mathbf{i})\ \Sigma(\mathbf{i}'=\mathbf{0},\mathbf{b})\ \Sigma(\mathbf{j}'=\mathbf{0},\mathbf{i}')\\ &\bullet\ M(\mathbf{i},\mathbf{j},\mathbf{a})\ M(\mathbf{i}',\mathbf{j}',\mathbf{b})\ f(\mathbf{i},\mathbf{j},\mathbf{a},\varphi_A,\theta_A)\ f(\mathbf{i}',\mathbf{j}',\mathbf{b},\varphi_B,\theta_B)\\ &\bullet\ W(A,\mathbf{g},n,\alpha)\ W(B,\mathbf{g}',m,\beta), \qquad (3.22)\end{aligned}$$

where the f and M coefficients and the **g** and **g**' vectors arise from the rotations that put each CETO in a well-oriented form respect the other. The W functions are WO-CETO's respect centers **B** and **A** respectively.

Collecting terms and using a compact form:

$$\begin{aligned}\Omega_{AB} &= \chi(\mathbf{A},\mathbf{a},n,\alpha)\ \chi(\mathbf{B},\mathbf{b},m,\beta)\\ &= \Sigma_i\ \Sigma_j\ C_i^{AB}\ C_j^{BA}\ W(\mathbf{A},\mathbf{a}_i,n,\alpha)\ W(\mathbf{B},\mathbf{b}_j,m,\beta)\\ &= \Sigma_\mu\ \Gamma_\mu\ \Omega_{AB,\mu}^{W}, \qquad (3.23)\end{aligned}$$

where the symbol $\Omega^{w}_{AB,\mu}$ means charge distribution between WO-CETO functions.

Then, in order to solve the CETO integral problem the general radial term:

$$P_{nm}(\mathbf{A},\alpha,\mathbf{B},\beta) = P_n(\mathbf{A},\alpha)\, P_m(\mathbf{B},\beta), \tag{3.24}$$

needs to be expanded in such a context where $x_{1A}=x_{1B}$, $x_{2A}=x_{2B}$ and the binomial expansion of $(x_{3B}+R)^n$ can be used to express x_{3A} in terms of x_{3B} and vice versa. Therefore, for example:

$$x^n_{3A} = (x_{3B}+R)^n = \sum_{i=0}^{n} \left\{ \binom{n}{i} R^{n-i} \right\} x^i_{3B}. \tag{3.25}$$

3.4. WO-CETO Products

WO-CETO products will be discussed here in all the related aspects, as a tool to construct integrals over CETO functions.

3.4.1. Cartesian Form in Terms of E-CETO's

The product of two WO-CETO functions can be written in terms of the third axis components, so:

$$\Omega^{w}_{AB} = W(A,\mathbf{a},n,\alpha)\, W(B,\mathbf{b},m,\beta)$$

$$= \Sigma(\mathbf{i}=\mathbf{0},\mathbf{p})\, \Gamma(\mathbf{i},\mathbf{a},\mathbf{b},R)\, x_1^{a_1+b_1}\, x_2^{a_2+b_2}\, x_3^{i_1+i_2}\, P_{nm}(\mathbf{A},\alpha,\mathbf{B},\beta), \qquad (3.26)$$

with $\Gamma(\mathbf{i},\mathbf{a},\mathbf{b},R) = (-1)^{a_3-i_1} \binom{a_3}{i_1} \binom{b_2}{i_2} (R/2)^{a_3+b_3-(i_1+i_2)}$.

The two dimensional vector **p** is defined as: (a_3,b_3) and being $P_{nm}(\mathbf{A},\alpha,\mathbf{B},\beta)$ defined in equation (3.24).

Finally, equation (3.26) show how CETO densities can be expressed as a combination of E-CETO densities:

$$V_{AB} = V(R/2,\mathbf{v},n,m,\alpha,\beta) = x_1^{v_1}\, x_2^{v_2}\, x_3^{v_3}\, P_{nm}(\mathbf{A},\alpha,\mathbf{B},\beta), \qquad (3.27)$$

where $n,m \in \{0,1\}$.

3.4.2. Prolate Spherical Coordinates Form

a) WO-CETO products as described in equation (3.26) can be expressed in terms of prolate spherical coordinates [42,47,53], and are defined here by means of the following relationships:

Definition 7: Prolate spherical coordinates.

1. Coordinates:

$$u=(r_A+r_B)\,R^{-1},\quad v=(r_A-r_B)\,R^{-1},\quad \varphi=\varphi_A=\varphi_B.$$

One can speak about this convention in terms of the **prolate spherical parameter system** $\{r_A,r_B,R\}$.

2. Volume element:

$$dV=(R/2)^3\,(u^2-v^2)\,du\,dv\,d\varphi \equiv (R/2)\,r_A\,r_B\,du\,dv\,d\varphi.$$

3. Integration limits: $u\in\{1,\infty\}$, $v\in\{-1,+1\}$, $\varphi\in\{0,2\pi\}$.

4. Useful expressions: Some formulae relating cartesian and spherical coordinate elements with prolate spherical variables are given here due to their later use.

4a. One can write for the vector radius and the trigonometric functions of the angle θ on both centers:

$$r_C=(R/2)(u+sv),\ \sin\theta_C = F(u,v)/r_C \text{ and}$$
$$\cos\theta_C = (1+suv)/r_C, \text{ with } C\in\{A,B\} \text{ and } s\in\{+1,-1\}.$$

4b. The cartesian coordinates can be also expressed as:

$$x_1 = (R/2)\, F(u,v)\, \cos\varphi,$$
$$x_2 = (R/2)\, F(u,v)\, \sin\varphi,$$
$$x_3 = (R/2)\, uv.$$

4c. Whenever used the function F(u,v) is here defined as:

$$F(u,v) = \{(u^2-1)\,(1-v^2)\}^{1/2}.\blacksquare$$

When dealing with integrals involving WO-CETO functions, as occurs within a STO framework, using a prolate spherical coordinate system like the described above is one of the best ways to obtain analytic formulae for two center molecular integrals.

b) At the moment of using such a prolate spherical coordinate system sometimes it is necessary to expand binomial expressions involving the prolate spherical coordinates {u,v}. This fact leads one to develop a technique in order to construct such binomial forms in terms of prolate spherical coordinate polynomials. Any of such binomial functions can be written generally as:

$$L(a,b,c,d,N) = (s_f\, u^a\, v^b + s_g\, u^c\, v^d)^N, \qquad (3.28)$$

where {a,b,c,d,N} are non-negative integers and $\{s_f, s_g\}$ plus or minus signs. Equation (3.28) can be rewritten, after expanding the binomial expression, as:

$$L(a,b,c,d,N) = (u^c\, v^d)^N \sum_{k=0}^{N} \Gamma(k,N)\, (u^{a-c}\, v^{b-d})^k, \tag{3.29}$$

where: $\Gamma(k,N) = \binom{N}{k} s_f^k\, s_g^{N-k}$.

Thus, a product of expressions like these of equation (3.28), that is:

$$M(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{N}) = \prod_k L(a_k,b_k,c_k,d_k,N_k), \tag{3.30}$$

can be written using equation (3.29) as:

$$M(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{N}) = u^{\mathbf{c}\bullet\mathbf{N}}\, v^{\mathbf{d}\bullet\mathbf{N}} \bullet \Sigma(\mathbf{k}=\mathbf{0},\mathbf{N})\, \Gamma(\mathbf{k},\mathbf{N})\, u^{(\mathbf{a}-\mathbf{c})\bullet\mathbf{k}}\, v^{(\mathbf{b}-\mathbf{d})\bullet\mathbf{k}}, \tag{3.31}$$

where the {u,v} coordinate exponents are dot products of the involved vectors, that is, for example: $\mathbf{c}\bullet\mathbf{N} = \Sigma_i c_i N_i$. The symbol $\Sigma(\mathbf{k}=\mathbf{0},\mathbf{N})$ corresponds to a nested summation and finally the coefficients $\Gamma(\mathbf{k},\mathbf{N})$ are in turn defined as the equation (3.29) coefficient products:

$$\Gamma(\mathbf{k},\mathbf{N}) = \prod_i \Gamma(k_i,N_i). \tag{3.32}$$

3.4.3. Elementary Integrals over Prolate Spherical Coordinates

a) The integrals over the u and v prolate spherical variables of the products M(**a**,**b**,**c**,**d**,**N**) multiplied by the exponential factor:

$$E(\sigma,\tau,u,v) = \exp(-\sigma u - \tau v), \tag{3.33}$$

where $\sigma>0$ and τ are real numbers, constitute a very general building block of the needed integral expressions over CETO functions. They can be written as:

$$Z(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{N},\sigma,\tau) = \Sigma(\mathbf{k}=\mathbf{0},\mathbf{N})\ \Gamma(\mathbf{k},\mathbf{N}) \cdot \iint u^{r(\mathbf{k})}\ v^{s(\mathbf{k})}\ E(\sigma,\tau,u,v)\ du\ dv, \tag{3.34}$$

where $r(\mathbf{k})=\mathbf{c}\cdot\mathbf{N}+(\mathbf{a}-\mathbf{c})\cdot\mathbf{k}\equiv r$ and $s(\mathbf{k})=\mathbf{d}\cdot\mathbf{N}+(\mathbf{b}-\mathbf{d})\cdot\mathbf{k}\equiv s$, thus the integral above can be expressed by means of the nested sum:

$$Z(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{N},\sigma,\tau) = \Sigma(\mathbf{k}=\mathbf{0},\mathbf{N})\ \Gamma(\mathbf{k},\mathbf{N})\ A_r(\sigma)\ B_s(\tau), \tag{3.35}$$

where the integral $A_r(\sigma)$ and $B_s(\tau)$ definitions can be found in section 3.4.4 below.

b) In particular, for some of the integrals that will appear later on, it is helpful to obtain, using the discussed general form in section

a) above, an expression for the expansion of the binomial products which can be written as:

$$P(\mathbf{a},u,v) = (u+v)^{a_1}\,(u-v)^{a_2}\,(1+uv)^{a_3}\,(1-uv)^{a_4}\,(u^2-1)^{a_5}\,(1-v^2)^{a_6}. \quad (3.36)$$

From the previous development one can easily deduce the equivalent compact form:

$$P(\mathbf{a},u,v) = \Sigma(\mathbf{r}=\mathbf{0},\mathbf{p})\;\Gamma(\mathbf{r},\mathbf{a})\;u^{r_1}\,v^{r_2}, \quad (3.37)$$

where $\mathbf{p}=(t+2a_5,t+2a_6)$ with $t=a_1+a_2+a_3+a_4$.

The expansion coefficients in equation (3.37) are defined as:

$$\Gamma(\mathbf{r},\mathbf{a}) = \Sigma(\mathbf{k}=\mathbf{0},\mathbf{a})\;\delta(k_1+k_2+k_3+k_4+2k_5=r_1) \cdot \delta(a_1+a_2-(k_1+k_2)+k_3+k_4+2k_6=r_2)\;C(\mathbf{k},\mathbf{a}), \quad (3.38)$$

where:

$$C(\mathbf{k},\mathbf{a}) = (-1)^q \prod_{i=1}^{6} \binom{a_i}{k_i}, \quad (3.39)$$

with the index definition: $q=a_2+a_5+k_2+k_4+k_5+k_6$.

A practical, convenient manner to obtain the coefficients $\Gamma(\mathbf{r},\mathbf{a})$ appearing in equation (3.37) consists of generating all the six involved sums, using a nested summation algorithm, employing a computational structure similar to the one shown in Program 1, adding the corre-

sponding contributions of $C(\mathbf{k},\mathbf{a})$ to the associated $\Gamma(\mathbf{r},\mathbf{a})$ coefficient determined by the values of r_1 and r_2, as shows schematically equation (3.38).

c) When these expansions in terms of prolate spherical coordinates are put in practice, equation (3.37) has the exponential function defined in equation (3.33), accompanying the expression of the most usual integrals needed in next sections. The product of $P(\mathbf{a},u,v)$ with $E(\sigma,\tau,u,v)$, upon integration leads to some simplified expressions where the integrals of type $A_k(\sigma)$ and $B_k(\tau)$, discussed next below in section 3.4.4, have a leading role:

$$Z(\mathbf{a},\sigma,\tau) = \int\int P(\mathbf{a},u,v)\, E(\sigma,\tau,u,v)\, du\, dv$$

$$= \Sigma(\mathbf{r}=\mathbf{0},\mathbf{p})\, \Gamma(\mathbf{r},\mathbf{a})\, A_{r_1}(\sigma)\, B_{r_2}(\tau). \qquad (3.40)$$

An integral like $Z(\mathbf{a},\sigma,\tau)$ can be evaluated collecting in a matrix the coefficients $\Gamma=\{\Gamma(\mathbf{r},\mathbf{a})\}$ and in a couple of column vectors the integrals $\mathbf{A}=\{A_{r_1}(\sigma)\}$ and $\mathbf{B}=\{B_{r_2}(\tau)\}$. Then one can write compactly:

$$Z(\mathbf{a},\sigma,\tau) = \mathbf{A}^T(\sigma)\, \Gamma(\mathbf{a})\, \mathbf{B}(\tau). \qquad (3.41)$$

d) Precisely, once the integrals $Z(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{N},\sigma,\tau)$ or $Z(\mathbf{a},\sigma,\tau)$ are evaluated, the one electron two center integrals over CETO functions, associated with a great deal of one electron operators, are easily computed. In the same manner as Wahl et al. [47] had shown, some

years ago, when working with STO functions. A discussion and various applications of this property will be performed in section 6 below.

3.4.4. Auxiliary Functions

When discussing the integrals over prolate spherical coordinates four kinds of auxiliary integrals appear. They will be studied by affine couples next.

3.4.4.1 $A_k(t)$ and $B_k(t)$ integral functions.

Due to the extreme importance throughout this paper it will be presented now a review of the two basic integral forms $A_k(t)$ and $B_k(t)$, with details on their computational properties and structure.

a) General Form.

The integrals $A_k(t)$ and $B_k(t)$, are well known auxiliary functions [10,42,50c,53]:

Definition 8.1: $A_k(t)$ integrals.

$$A_k(t) = \int_1^{\infty} x^k\, e^{-tx}\, dx = \{\, k!e^{-t}/t^{k+1} \,\} \sum_{p=0}^{k} t^p/p!.\blacksquare \qquad (3.42)$$

For $A_k(t)$ a well known upwards recurrence formula is used [42,53]:

$$A_k(t) = (k\, A_{k-1}(t) + \exp(-t))/t. \qquad (3.43)$$

Definition 8.2: $B_k(t)$ integrals.

$$B_k(t) = \int_{-1}^{+1} x^k e^{-tx}\, dx = \{ k!/t^{k+1} \} \{ \sum_{p=0}^{k} ((-1)^p e^t - e^{-t})\, t^p/p! \}, \qquad (3.44)$$

with $B_k(0) = 2\delta(k=\dot{2})/(k+1)$.■

Although for the $B_k(t)$ functions an upwards recurrence formula exists:

$$B_k(t) = ((-1)^k \exp(t) - \exp(-t) + nB_{k-1}(t))/t, \qquad (3.45)$$

the $B_k(t)$ integrals are not so accurately obtained due to the sign alternation in the expressions (3.44) or (3.45) and the possible different magnitudes of e^{-t} and e^t, given a fixed value of the argument t. Trying to overcome this drawback, it has been used here the Taylor expansion of the inverse exponential function on the $B_k(t)$ definition, in the same manner as Zener and Guillemin [54a] early in 1929 or later on De Jeu [54b] and Saunders [42] recommend. Upon Taylor expansion integration a new $B_k(t)$ expression results:

$$B_k(t) = 2 \sum_{p=0}^{\infty} \delta(p+k=\dot{2})\, (-t)^p / \{ p!\, (p+k+1) \}. \qquad (3.46)$$

The analytic form of $B_k(t)$ in equation (3.46) contains a power series, but convergence is fast due to the presence of the inverse factorial and to the fact that k and p must have the same parity, overriding sign alternation.

Evaluation of $B_k(t)$ can always be obtained with positive arguments:

$$B_k(-t) = (-1)^k B_k(t), \tag{3.47}$$

and for even and odd k values the following formula can be used, after rearranging equation (3.46):

$$B_k(t) = 2c \sum_{q=0}^{\infty} t^{2q} / \{ (2q+a)! (2(q+b)+1) \}, \tag{3.48}$$

where $a=\delta(k\neq 2)$, $b=[\![k/2]\!]+a$ and $c=(-t)^a$.

3.4.4.1. Products of the Integrals A and B

a) In some places it is needed the integral product where both integral factors have the same arguments: $A_r(t)B_s(t)$, in the usual applications as shows section 3.4.2 above, and sections 6 and 7. From Definitions 8.1 and 8.2, one can write:

$$A_r(t)B_s(t) = t^{-(r+s+2)} \sum_{p=0}^{r} \sum_{q=0}^{s} \Lambda(w) \{ (-1)^q - e^{-2t} \} t^{p+q}$$

$$= \sum_{L=0}^{M} \{ \Gamma_L(-1) - \Gamma_L(+1) e^{-2t} \} t^{-(L+2)}, \tag{3.49}$$

where M=r+s and $\Lambda(\mathbf{w})=(w_1!w_2!)/(w_3!w_4!)$ being in this case $(w_1,w_2,w_3,w_4)=(r,s,p,q)$. If $\sigma\in\{-1,+1\}$ is a sign then the coefficients $\Gamma_L(\sigma)$ are defined as:

$$\Gamma_L(\sigma) = \sum_{p=0}^{r} \sum_{q=0}^{s} \delta(p+q-(r+s)=L)\, \Lambda(\mathbf{w})\, \sigma^p. \qquad (3.50)$$

b) The same product as before but with different arguments is also needed in some integral calculations. Thus, it may be interesting to develop its possible form. The final result is:

$$A_r(t)B_s(u) = \{ r!s!/(t^{r+1}\, u^{s+1}) \} \sum_{i=0}^{r} \sum_{j=0}^{s} \{ (-1)^j\, e^{u-t} - e^{-(u+t)} \}/\{i!j!\}\, t^i u^j, \qquad (3.51)$$

or, using the form (3.48) for $B_s(u)$:

$$A_r(t)B_s(u) = 2(-u)^a\, e^{-t} \sum_{p=0}^{r} \sum_{q=0}^{\infty} r!/\{p!(2q+a)!(2(q+b)+1)\}\, t^{p-(r+1)}\, u^{2q}\,, \qquad (3.52)$$

being $a=\delta(s\neq\dot{2})$ and $b=[\![s/2]\!]+a$.

3.4.4.2. $C_k(t)$ and $D_k(t)$ Integral Functions

a) When dealing with Coulomb operators and in some other cases, which will be discussed in sections 6 and 7, one needs another kind of integral related to the $A_k(t)$ functions defined previously in equation (3.42).

These integrals can be described as follows:

Definition 8.3: $C_k(t)$ integrals.

$$C_k(t) = \int_1^{\infty} x^{-k}\, e^{-tx}\, dx, \blacksquare \qquad (3.53)$$

being $C_k(t)$ a member of the family of the so called exponential integrals. The evaluation and computational properties of these integrals can be extensively found in the handbook of Abramowitz and Stegun [50c].

The most interesting features will be given here, in order to provide the potential reader with the necessary information as it has been done when the integrals A and B were studied before. It is not surprising that the integral (3.53) as well as the parent A integrals can be related to the well known incomplete gamma function conspicuously present in the GTO Coulomb type integrals [42,50c].

A series expansion or a continued fraction can be used to compute $C_k(t)$. Also a recurrence relation is easily found:

$$C_{k+1}(t) = (e^{-t} - t\, C_k(t))/k. \qquad (3.54)$$

The $C_k(t)$ asymptotic values are also well defined. One has for example:

$$C_k(0) = (k-1)^{-1},\ k \neq 1, \qquad (3.55)$$

and in reference [50c] expressions for large k and t values can also be found.

b) Still another integral will be also needed, this time connected with the $B_k(t)$ integrals (3.45). Their definition is:

Definition 8.4: $D_k(t)$ integral.

$$D_k(t) = \int_0^1 x^k \, e^{-tx} \, dx, \blacksquare \tag{3.56}$$

which is easily connected by a trivial change of variable to the incomplete gamma function [42,50c].

Some characteristics of the integrals (3.56) have been discussed early by O-ohata and Ruedenberg [26b], where the recurrence relation is described:

$$D_k(t) = (k \, D_{k-1}(t) - t \, e^{-t}) / t, \tag{3.57}$$

which is well behaved in downward direction. From the same reference a power series can be used:

$$D_k(t) = e^{-t} \sum_{p=1}^{\infty} t^{p-1} / [(k+1)(k+2)...(k+p)]. \tag{3.58}$$

Extreme parameter values can also be easily computed:

$$D_0(t) = (1 - e^{-t})/t, \tag{3.59}$$

and also one has:

$$D_k(0) = (1+k)^{-1}. \tag{3.60}$$

A simple change of variable in the $D_k(t)$ integral permits to express it as a finite linear combination of $B_k(t)$ integral functions.

4. Two Center Expansion of a CETO Function Product

As it is well known, see, for example, references [17b,47], a good starting point of AO integral construction resides not only in the proper ETO functions as such, but in the form of ETO products themselves.

This consideration is reinforced by the fact that here is described, from a practical point of view, after the theoretical development of section 3, how to optimally express a CETO product in terms of a linear combination of CETO's, centered at each of the sites belonging to the product functions.

A short discussion of the meaning and the possible utilization of this contingency are also present at the end of this section.

4.1. General Outline

a) Suppose one has two CETO functions centered at some points {**A**,**B**} at a distance R: $\{\chi_A,\chi_B\}$. The product of both functions Ω_{AB} as in equation (3.22) has a relevant part:

$$\varepsilon_{AB} = P_{00}(\mathbf{A},\alpha,\mathbf{B},\beta) = \exp(-\alpha r_A)\,\exp(-\beta r_B), \tag{4.1}$$

and a marginal one: the rest of the product.

b) In the present approach Ω_{AB} will be expanded, and at the same time ε_{AB}, in terms of CETO's centered separately on **A** and **B**, using a more general structure than the one proposed by Ruedenberg [38], consisting in the sum of two bilinear forms:

$$\Omega_{AB} \approx \sum_{\alpha\in A}\sum_{\alpha'\in A} \Gamma_\alpha\,\Gamma_{\alpha'}\,\Omega_{\alpha\alpha'} + \sum_{\beta\in B}\sum_{\beta'\in B} \Gamma_\beta\,\Gamma_{\beta'}\,\Omega_{\beta\beta'}. \tag{4.2}$$

But as it has been shown before in equation (3.21) the CETO products $\Omega_{\alpha\alpha'}$ and $\Omega_{\beta\beta'}$ can be substituted by unique CETO functions χ_a and χ_b, respectively. When terms are collected in equation (4.2), taking into account the discussed one center CETO property it is obtained:

$$\Omega_{AB} \approx \sum_{a\in A} g_a\,\chi_a + \sum_{b\in B} g_b\,\chi_b. \tag{4.3}$$

In fact, after obtaining this simplified result, it is crucial to prove the exponential product ε_{AB} can be expanded at the centers **A** and **B** with sufficient accuracy to be considered a good starting point, before studying the general case represented by the CETO product in equation (3.22). A detailed analysis will be performed next in section 5, after the general theoretical discussion on the way these linear combinations can be calculated, which follows.

c) According to the previous discussion the product Ω_{AB} can be also expressed in terms of WO-CETO functions $\{\psi_\alpha^A,\psi_\beta^B\}$ centered separately on **A** and **B**. Connecting the present approach with old ideas of Harris and Rein [55a], Newton [55b] and Okninski [55c], one can write:

$$\Omega_{AB} \approx \sum_{\alpha\in A} c_\alpha^A \psi_\alpha^A + \sum_{\beta\in B} c_\beta^B \psi_\beta^B. \tag{4.4}$$

Optimal linear coefficients can be computed using the compact matrix form of equation (4.4), supposing both sides equal:

$$\Omega_{AB} = \psi^A \mathbf{c}^A + \psi^B \mathbf{c}^B, \tag{4.5}$$

then $\{\mathbf{c}^A,\mathbf{c}^B\}$, the expansion coefficients column vectors, can be obtained by multiplying on the left equation (4.5) by the WO-CETO basis set column vectors $(\psi^A)^+$ and $(\psi^B)^+$ where both basis sets are stored respectively, and integrating. The following linear system can be built up:

$$\begin{aligned} \mathbf{s}_A &= \mathbf{S}_{AA}\,\mathbf{c}^A + \mathbf{S}_{AB}\,\mathbf{c}^B, \\ \mathbf{s}_B &= \mathbf{S}_{BA}\,\mathbf{c}^A + \mathbf{S}_{BB}\,\mathbf{c}^B, \end{aligned} \tag{4.6}$$

where $\mathbf{s}_A=\langle\psi^A \mid \Omega_{AB}\rangle$ and $\mathbf{S}_{AA}=\langle\psi^A \mid \psi^A\rangle$ is the metric matrix of the WO-CETO basis set centered on **A**. The same definitions apply but with respect of center **B** for $\mathbf{s}_B$ and $\mathbf{S}_{BB}$.

d) Collecting $\{\mathbf{s}_A,\mathbf{s}_B\}$ and the other matrices in (4.6) in a hypermatrix system one can easily write:

$$\mathbf{s} = \mathbf{S}\,\mathbf{c}, \tag{4.7}$$

with $\mathbf{s} = \begin{pmatrix} \mathbf{s}_A \\ \mathbf{s}_B \end{pmatrix}$ and obvious similar definitions for the organization of the remaining hypermatrices in the equation (4.7). As **S** is the metric matrix of the union of the basis sets $\{\psi^A,\psi^B\}$, and **S**>0 consequently, then an inverse will always exist: $\mathbf{S}^{-1}$, which can be used to solve the equation (4.7):

$$\mathbf{c} = \mathbf{S}^{-1}\,\mathbf{s}. \tag{4.8}$$

4.2. General Linear Parameter Optimization Procedure

Cholesky's decomposition [56a,b] has been used to obtain $\mathbf{S}^{-1}$ in a practical manner. The reason to choose Cholesky's algorithm is found essentially in the numerical stability of this procedure, but alternatively in the possible definition of a recursive pathway to evaluate $\mathbf{S}^{-1}$ [56c], starting from a small number of functions up to any

bigger dimension adding a function element or a functional subset of them each time as follows:

Procedure 2: Inverse of a Definite Positive Matrix using Cholesky's Decomposition.

a) Suppose known $\mathbf{S}_n$ at some dimension n of the functional subspace.

b) Cholesky's decomposition of $\mathbf{S}_n$ is defined by means of an upper triangular matrix $\mathbf{T}_n$ and the transpose $\mathbf{T}_n^T$:

$$\mathbf{S}_n = \mathbf{T}_n^T \mathbf{T}_n. \tag{4.9}$$

then the inverse of $\mathbf{T}_n$, $\mathbf{T}_n^{-1}$ is easily computed [56d] and:

c) $$\mathbf{S}_n^{-1} = \mathbf{T}_n^{-1} (\mathbf{T}_n^{-1})^T. \blacksquare \tag{4.10}$$

There is no need to repeat the same pathway again if $\mathbf{S}_n$, $\mathbf{T}_n$ and $\mathbf{T}_n^{-1}$ as well as $\mathbf{S}_n^{-1}$ are known at some dimension level n. The following recursive algorithm has no difficulty whatsoever:

Procedure 3: Recursive Inverse Calculation using Cholesky's Decomposition

a) Suppose the (n×m) matrix $\mathbf{s}_m$ and the (m×m) matrix σ_m are known and the structure of the metric at the dimension level n+m can be written as:

$$\mathbf{S}_{n+m} = \begin{bmatrix} \mathbf{S}_n & \mathbf{s}_m \\ \mathbf{s}_m^T & \sigma_m \end{bmatrix}. \tag{4.11}$$

b) Assuming the following structure for Cholesky's decomposition of a triangular matrix at dimension level n+m is:

$$\mathbf{T}_{n+m} = \begin{bmatrix} \mathbf{T}_n & \mathbf{t}_m \\ \mathbf{0}^T & \tau_m \end{bmatrix}, \tag{4.12}$$

then, it is easy to see that:

$$\mathbf{t}_m = (\mathbf{T}_n^{-1})^T \ \mathbf{s}_m, \tag{4.13}$$

and

$$\tau_m^T \tau_m = \sigma_m - \mathbf{t}_m^T \mathbf{t}_m, \tag{4.14}$$

that is: τ_m corresponds to the Cholesky's decomposition upper triangular matrix of the (m×m) matrix appearing on the right hand side of equation (4.14).

c) There is also needed the matrix $\mathbf{T}_{n+m}^{-1}$, which can be obtained with:

$$\mathbf{T}_{n+m}^{-1} = \begin{bmatrix} \mathbf{T}_n^{-1} & \mathbf{t}_m^{(-1)} \\ \mathbf{0}^T & \tau_m \end{bmatrix}, \tag{4.15}$$

in this manner the (m×n) matrix $\mathbf{t}_m^{(-1)}$ must be defined as:

$$\mathbf{t}_m^{(-1)} = - \mathbf{T}_n^{-1} \mathbf{t}_m \tau_m^{-1}. \tag{4.16}$$

d) And the new inverse is simply:

$$\mathbf{S}_{n+m}^{-1} = \begin{bmatrix} \mathbf{S}_n^{-1} + \mathbf{t}_m^{(-1)} (\mathbf{t}_m^{(-1)})^T & \mathbf{t}_m^{(-1)} (\tau_m^{-1})^T \\ \tau_m^{-1} (\mathbf{t}_m^{(-1)})^T & \tau_m^{-1} (\tau_m^{-1})^T \end{bmatrix}. \blacksquare \tag{4.17}$$

Thus, the described recursion above needs only the (n×m) matrices: **s**, **t**, $\mathbf{t}^{(-1)}$ and the square (m×m) matrices τ and τ^{-1} at each step.

Procedures 2 and 3 are very convenient when looking for a quick test on the efficiency of adding or not a new CETO set to the expansion (4.4), while optimizing the non-linear parameters. All the procedure becomes simpler when only one function is added, because in this case one has m=1. The (n×m) matrices appear to be column vectors (n×1) and the square (m×m) σ and τ matrices become (1×1), scalars: that is σ=1 and τ=$(1 - \mathbf{t}_n^T \mathbf{t}_n)^{1/2}$.

4.3. Final Remarks

4.3.1. Least Squares Fit

Moreover, this computational scheme is the same as applying a least squares fit to the expression (4.4) or (4.5). That is, solution of the system (4.7) minimizes the quadratic error $q^{(2)}$ with respect to the linear parameters $\{c^A,c^B\}$:

$$q^{(2)} = | \Omega_{AB} - (\psi^A c^A+\psi^B c^B) |^2. \tag{4.18}$$

Supposing normalization of all the involved functions, then the measure:

$$\xi = \int \Omega_{AB} (\psi^A c^A+\psi^B c^B) \, dV, \tag{4.19}$$

is nothing more than the degree of fitting of the linear combination (4.5) of the WO-CETO charge distribution Ω_{AB}: as $\xi \rightarrow 1$ better Ω_{AB} is covered by the two center expansion and $q^{(2)} \rightarrow 0$.

Moreover, one can think of adjusting the linear combination (4.5) to the charge distribution Ω_{AB} using as a weight some positive definite operator Θ. In this case, one has to optimize the integral:

$$\xi' = \int \Omega_{AB} \, \Theta \, (\psi^A c^A+\psi^B c^B) \, dV. \tag{4.20}$$

Here, the first option was adopted. But when dealing with Coulomb integrals using of the expression (4.20) could lead to more accurate results if Θ is replaced by the related operator.

4.3.2. Reduction of Three and Four Center Integrals to Mono and Bicentric Terms

Using the CETO properties outlined in sections 3 and the present one, while considering expression (4.4), it can be seen that molecular integrals involving charge distributions may be constructed as linear combinations of other integrals within a reduced number of centers.

For example, three center nuclear attraction integrals (AB | C) will reduce to an expression involving two center integrals of (AA | C) and (BB | C) type.

Also, Coulomb repulsion integrals will be simplified to a combination of two electron integrals like:

$$\langle \psi_A | r_{12}^{-1} | \psi_B \rangle = \iint \psi_A(1)\, \psi_B(2)\, r_{12}^{-1}\, dV_1\, dV_2. \qquad (4.21)$$

To illustrate this possibility, let us consider a two electron four center integral of any kind: (AB:CD). Considering and expanding both involved charge distributions as described in the expression (4.4), then the integral will be a sum of four bilinear terms which will depend on the four integral kinds with a structure similar to equation (4.21), that is: (A:C), (A:D), (B:C) and (B:D). This will be further analyzed in sections 7.3 and 7.4.

5. Expansion of the Product of Two Exponential Functions Centered at A and B

The previous general discussion carried out in the sections above, comes to a point where the relevant exponential part (4.1) of the CETO product (3.22) must be studied in depth.

A possible computation of three and four center integrals rely upon the development of the exponential product covering by means of chosen WO-CETO basis set linear combinations. The integrals and techniques needed to optimize the expansion functions will be described here.

5.1. Basis Sets

To prove the viability of the procedure outlined previously in section 4, the problem to expand expression ε_{AB} as in equation (4.1) in terms of the product of two 1s STO or CETO functions, will be considered. This simple form involving even E-CETO's has been chosen as a check of the preceding framework and because of the arguments discussed in sections 3.2.c and .d above.

a) A particular form of the expansion given in equation (4.5) will be used here, because each term of the chosen basis set may be written as a linear combination of two WO-CETO functions:

$$\psi_i = x_{3A}^{a_i} \exp(-\alpha_i r_A) + f\, x_{3B}^{b_i} \exp(-\beta_i r_B), \tag{5.1}$$

and cast in the expression (4.1) of ε_{AB}, like:

$$\varepsilon_{AB} \approx L_n = \sum_{i=1}^{n} c_i \psi_i. \tag{5.2}$$

b) The following constraints were also imposed when defining the functions described in equation (5.1):

1. α_i, β_i are real positive parameters to be optimized, obeying the ratio: $\alpha / \beta = \alpha_i / \beta_i$, $\forall i$.
2. f is a prefixed factor used to correct the expression of equation (5.1) when $\alpha \neq \beta$ and it is computed with the follo-

wing recipe. If R is the intercenter distance and calling $e_A = e^{-\alpha R}$ and $e_B = e^{-\beta R}$, then the factor f can be defined as:

$$f = e_A (e_B - 1) / \{ e_B (e_A - 1) \}. \qquad (5.3)$$

This takes into account, when $r_A=r_B=0$, that the ratio of ε_{AB} factors in equation (4.1), is preserved into equation (5.2).

3. $a_i=b_i$ are non-negative integers to be optimized.

c) Finally, one must be aware of the fact that when using equation (5.1) into equation (4.4) and after adopting the mentioned above constraints **1.** to **3.**, this situation forces the set ψ^A to determine the set ψ^B.

Of course, other choices can be made at the moment to define the ψ functions, demanding a greater non-linear parameter optimization work. For example, one can propose the general basis set element form:

$$\psi_i = P_{m_i}(x_{3A})\, Q_{n_i}(r_A) \exp(-\alpha_i r_A) + P_{t_i}(x_{3B})\, Q_{u_i}(r_B) \exp(-\beta_i r_B), \qquad (5.4)$$

where $P_k(x)$ and $Q_k(x)$ are arbitrary polynomials of degree k, but in this preliminary stage these possible options have been discarded, using Occam's razor, for the simplest one.

5.2. Integrals Involved in the ε_{AB} Expression

The integrals entering two center expansions parameter calculation as given in equation (5.1) will be discussed here.

The basis functions represented in equation (5.1) have to be optimized along the linear combination described in equation (5.2). The whole optimization process needs the following integrals in order to be implemented:

a) A two center integral defined as a simplified form of equation (3.40):

$$\mathcal{E}(a,\alpha,b,\beta,R) = \int x_{3A}^{a}\ x_{3B}^{b} \exp(-\alpha r_A - \beta r_B)\, dV, \qquad (5.5)$$

which can be calculated using a prolate spherical coordinate system as described in Definition 7, and employing the simplification:

$$P(\mathbf{a},b,u,v) = (u+v)\,(u-v)\,(1+uv)^{a}\,(1-uv)^{b} = P(\mathbf{c},u,v)$$

$$= \Sigma(\mathbf{r}=\mathbf{0},\mathbf{t})\ \Gamma(\mathbf{r},\mathbf{c})\ u^{r_1}\ v^{r_2}, \qquad (5.6)$$

where $\mathbf{t}=(t,t)$ with $t=a+b+2$, as it can be deduced from the general development discussed in section 3.4.2.b. The function $P(\mathbf{c},u,v)$ has been taken from equation (3.36) with the hexadimensional vector $\mathbf{c}=(1,1,a,b,0,0)$. With this previous information the integral (5.5) can be written as:

$$\mathcal{E}(a,\alpha,b,\beta,R) = 2\pi(R/2)^{t+1}\,\Sigma(r=0,t)\,\Gamma(r,c)\,A_r(\sigma)\,B_s(\tau), \qquad (5.7)$$

where:

$$\sigma=R(\alpha+\beta)/2 \quad \text{and} \quad \tau=R(\alpha-\beta)/2. \qquad (5.8)$$

b) A monocentric integral is also needed, defined as:

$$\mathcal{F}(a,\alpha) = \int z^a\, e^{-\alpha r}\, dV = \delta(a=\dot{2})\, 4\pi\, (a+2)!\,/\,\{(a+1)\,\alpha^{a+3}\}. \qquad (5.9)$$

c) Starting from the above integral definitions it is easy to compute the overlap integral between the exponential product ε_{AB} and the expansion basis set functions (5.1):

$$\langle\psi_i \mid \varepsilon_{AB}\rangle = N_\varepsilon N_i\, \{\mathcal{E}(a_i,\alpha+\alpha_i,0,\beta,R) + f\mathcal{E}(0,\alpha,b_i,\beta+\beta_i,R)\}, \qquad (5.10)$$

and the metric matrix element corresponding to the expansion basis set (5.1):

$$\begin{aligned}\langle\psi_i \mid \psi_j\rangle = N_i N_j\, \{\, &\mathcal{F}(a_i+a_j,\alpha_i+\alpha_j) + \mathcal{F}(b_i+b_j,\beta_i+\beta_j) \\ &+ f\,(\,\mathcal{E}(a_i,\alpha_i,b_j,\beta_j,R) + \mathcal{E}(b_i,\beta_i,a_j,\alpha_j,R)\,)\,\}. \end{aligned} \qquad (5.11)$$

d) The normalization factors $\{N_\varepsilon,N_i\}$ are also evaluated utilizing the previously defined integrals in a) and b) above:

$$N_i^{-2} = \mathcal{F}(2a_i,2\alpha_i) + f^2\,\mathcal{F}(2b_i,2\beta_i) + 2\mathcal{E}(a_i,\alpha_i,b_i,\beta_i,R), \qquad (5.12)$$

and

$$N_\varepsilon^{-2} = \mathcal{E}(0,2\alpha,0,2\beta,R). \qquad (5.13)$$

5.3. Scaling

a) Another very convenient property, which has been used in practice here at the moment to obtain numerical results, is based on the invariance of some CETO integral expressions upon scaling.

A more general but similar scaling property has been noticed by O-ohata, Taketa and Huzinaga [57] when dealing with GTO basis sets. It is easy to see that WO-CETO function exponents are related to the scaling of the A-B intercenter distance.

A change of scale on the $\mathcal{E}$ and $\mathcal{F}$ integrals, as defined in equations (5.5) and (5.9), can be made without effort using the substitutions:

$$R \rightarrow R^* = \zeta R \quad \text{and} \quad \alpha \rightarrow \alpha^* = \alpha/\zeta, \qquad (5.14)$$

where ζ is a scale factor and the star denotes scaled variables, functions or integrals.

b) It can be shown that:

$$\mathscr{E}^*(a,\alpha^*,b,\beta^*,R^*) = \mathscr{E}(a,\alpha/\zeta,b,\beta/\zeta,\zeta R) = \zeta^{a+b+3}\,\mathscr{E}(a,\alpha,b,\beta,R). \quad (5.15)$$

c) A similar relationship can be found among the $\mathscr{S}$ integrals:

$$\mathscr{S}^*(a,\alpha^*) = \delta(a=\dot{2})\,4\pi\,(a+2)!\,/\,((a+1)\,(\alpha/\zeta)^{a+3}) = \zeta^{a+3}\,\mathscr{S}(a,\alpha). \quad (5.16)$$

d) Scaling can be furthermore transferred to the normalization factors, as:

$$N_\varepsilon = \zeta^{(a+b+3)/2}\,N_\varepsilon^*, \quad (5.17)$$

and

$$N_i^* = (\mathscr{S}_a^* + \mathscr{S}_b^* + 2\mathscr{E}^*)^{-1/2} = (\zeta^{2a+3}\mathscr{S}_a + \zeta^{2b+3}\mathscr{S}_b + 2\zeta^{a+b+3}\mathscr{E})^{-1/2}. \quad (5.18)$$

e) The most useful aspect of this property appears when looking at the optimization of WO-CETO products related by scaling. As it can be easily deduced when using $a_i=b_i$ in equation (5.1), the following relationship holds:

$$\langle\psi_i^*\,|\,\varepsilon_{AB}^*\rangle = \langle\psi_i\,|\,\varepsilon_{AB}\rangle,\ \forall i, \quad (5.19)$$

consequently the solutions of the system (4.7) are invariant.

f) Even if two WO-CETO expansions cannot be directly related by scaling, a similar, already computed and optimized, WO-CETO expression can be chosen as a starting point in order to optimize non-linear parameters present in equation (4.7).

A possible strategy will be described by the following scheme:

Procedure 4: **Scaling**

1. Compute a scale factor ζ such as to have $R^*=2$. One can make some simplifications choosing $R^*=2$ due to the fact that the term R/2 appears in many expressions when dealing with the kind of integrals studied here.
2. Choose as optimization starting point the nearest optimal known expression of type (5.2) to the described exponential product in equation (4.1).
3. Optimize the actual expansion.
4. Recover the initial normalization factors and exponents.■

5.4. Non-linear Parameter Optimization

So far only the function forms and the optimization of linear parameters have been discussed, but the WO-CETO chosen basis set in order to expand the exponential product (4.1), even in the simplest form used here, demands optimization of some nonlinear parameters. In this subsection, the related problems are described and discussed.

5.4.1. General Considerations

Given a WO-CETO product, the goal is to expand the density function exponential part, ε_{AB}, in equation (5.2) as a linear combination I_{l_n} of some n functions having several non-linear degrees of freedom, as these in equation (5.1) have, which can be optimized in order to obtain the best fit of one to another.

Within the approach used here, and following the structure of the functions in equation (5.1), the variational non-linear parameters are the integer exponents of the x_3 variable: $\{a_i,b_i\}$ and the real positive defined exponential scale factors: $\{\alpha_i,\beta_i\}$.

Moreover, according to the constraints imposed here on the basis functions (5.1), one has always $a_i=b_i$ and the ratio α_i/β_i is maintained constant for every pair of functions. Thus, in this framework only two non-linear parameters must be considered for each basis function: {p,e}, the variable x_3 integer exponent p and the exponential real scale factor e, say.

A possible recursive algorithm, adapted to the already described Cholesky Procedures 2 and 3, in order to build up the coefficients of the linear combination (5.2) can be described as in section 4.2 above. In this preliminary stage the variable m involved in Procedure 3 is taken as unity, that is: only one function is added at each computational step. But, in general, m functions can be added at any time and in this manner has been the Cholesky algorithm described.

5.4.2. Non-linear Parameter Optimization Procedure Description: One function at a Time

The optimization of basis set non-linear parameters, appearing in equation (5.2), constitute one of the main steps in the preliminary work before many center integral evaluation. There will be described only a step by step procedure in order to optimize non-linear parameters of the involved functions one by one.

Procedure 5: **Non-linear Optimization**

Supposing known the optimal linear coefficients of a linear combination such as the one appearing in equation (5.2) and composed of n terms, L_n, the search for a new optimal function ψ_{n+1} in order to construct L_{n+1} is performed in the following steps:

1. Sweep a given range of {p,e} values in order to find some approximate starting point: $\{p_0,e_0\}$.
2. A bisection search is applied over a given interval around e_0 in order to ameliorate the covering of ε_{AB} by the linear combination L_{n+1}.
3. For every pair {p,e} tested in steps 1. and 2. above it is necessary to solve the linear system outlined in equation (4.7). When the linear coefficients are computed as in (4.8), then the overlap ξ, defined as in equation (4.19), between ε_{AB} and L_{n+1} can be evaluated.

4. As the name of the game is to optimize the objective function ξ, which signals the degree of similarity between the exponential product ε_{AB} and the linear combination L_{n+1}, then Cholesky's algorithm as described in Procedures 2 and 3 is very convenient for this purpose. This can be easily seen because when the non-linear parameters {e,p} change, only one column and the corresponding row of the basis set overlap matrix **S** in expression (4.7) change. In order to construct a smooth algorithm which can be adapted to the optimization steps as well as possible, it is necessary to organize the overlap matrix **S** to have the (n+1)-th test function associated to the last matrix column and row.
5. When the optimal (n+1)-th function is found, one can scan the previous n functions, re-optimizing their exponential scale factors. Several sweeps may be performed until the desired accuracy is reached.
6. Optionally the x_3 exponents can be re-optimized too. However, this integer optimization seems very insensitive to the addition of new functions. The best strategy consists in testing if some possible variation of the exponents as well as the attached exponential scale factors optimizes the overlap ξ, once all the exponential scale factors had been previously set and refined.■

5.4.3. Some Remarks on Optimization

Other optimization paths are possible. One can think that a covering of the exponential product ε_{AB} may be better achieved by a simultaneous n-dimensional optimization process like Newton-Raphson procedure [58]. In fact, however, the overlap ξ as described in equation (4.19) is an irregular function of the {p,e} parameter pairs, and possesses multiple maxima and minima with respect to the scale factors variation. A method with a scanning or sweep feature among a given parameter range will be perhaps superior in this case than a global search procedure, and so was adopted.

Parallel computations can be also imagined and used in parts of the optimization algorithm.

The characteristics of Cholesky's procedure allow one to construct computational codes working in parallel form, [52a,b], testing the effect of several functions on the linear coefficients concurrently.

Another alternative choice, this time referring to non-linear parameters, may consist in using a parallel algorithm to divide each exponent sweep into several regions, which can be analyzed simultaneously.

Also, the same can be said when a Simplex [58] algorithm framework is used: n+1 directions can be tested separately on the non-linear parameters in this case, each one running independently on a separate CPU.

These alternative possibilities will be taken into account and explored in the future.

5.5. Computational Example

Various numerical results and the computational details, related with practical case parameter evaluation in equations (5.1) and (5.2) are described in this section.

5.5.1. Hardware and Software Details

All calculations presented in this paper have been performed in AT 386 compatible clonic PC machines. Lahey [59], Microway NDP [60] and Watcom [61] Fortran compilers, working in a MS DOS [62] operating system environment, have been extensively used.

This proves the possible development of new computing ideas [63] within the expanding galaxy of inexpensive hardware and software: see [64] as an example of a detailed catalog of available

programs, as a test of this eventuality. Other authors like Tai [65] have been using this kind of hardware to study similar problems.

Source codes built up when constructing this paper can be transferred to usual computers (or supercomputers) without noticeable changes, following a described programming philosophy and discipline [66]. They will be published elsewhere.

Plans to test the extent to which parallel calculations can be done over CETO functions, using appropriate algorithms implemented on concurrent devices, like transputer boards [51b,c] connected to PC machines, are under way.

5.5.2. Covering of ε_{AB} Function

Some values of the overlap ξ expressed as in equation (4.19) are shown in **Table 5.1**. One can expect that when an overlap tends more and more to unity better should be the covering results concerned with the involved expansion.

Thus, looking at **Table 5.1** one can see how a convenient limited number of functions appear to be sufficient to attain high precision convergence.

Table 5.1

Variation of the Overlap $\langle L_n | \varepsilon_{AB} \rangle$ for R=2 and $\alpha=\beta=1$

Number of Functions n	Relative Optimization Time	$\langle L_n \| \varepsilon_{AB} \rangle$ Overlap
1	1	0.94149675
2	22	0.99549951
3	75	0.99919639
4	170	0.99951238
5	412	0.99975461
6	483	0.99993654
7	695	0.99996040
8	956	0.99997370
9	992	0.99998439
10	1735	0.99999256
11	2264	0.99999591
12	2899	0.99999902
13	3723	0.99999962
14	4679	0.99999985
15	5850	0.99999992

In order to have a visualization of the degree of covering of the ε_{AB} function in terms of the relative error committed when developing the proposed approach, **Figure 5.1** shows the error between ε_{AB} values and a linear combination: L_{20}, defined using equation (5.2), after complete linear and non-linear parameter optimization.

The error on plane X=0 is represented as well as the companion planes: X=2, X=4 and X=6. The intercenter distance is R=2 and the two exponents chosen as: $\alpha=\beta=1$.

One can see, when inspecting **Figure 5.1**, how the covering error is almost negligible when taking into account the involved

function values. The error raises in the measure that the space zone where the function vanishes is considered for analysis.

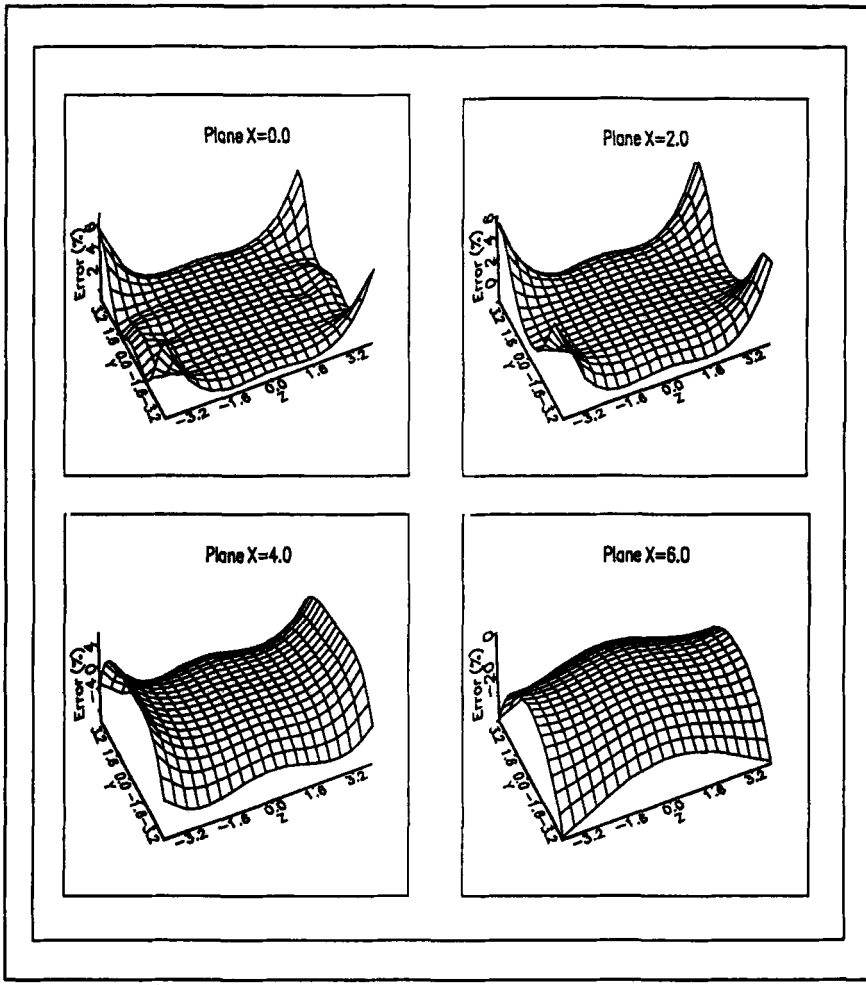


Figure 5.1: Error for L_{20} over several assorted planes.

6. Integrals over One-electron Operators

Although already described by Saunders [42], overlap and kinetic energy integrals over a CETO framework will be briefly discussed for completeness sake and at the same time giving more details, extending the information as far as possible and adding the expression of the variation of the electron mass with velocity and electric moment integrals, which can be considered as some sort of overlap integral extension.

Nuclear Attraction integrals are also studied at this section end, providing the first application example of the technique outlined in section 5 above.

Also the most important point has been taken into account here: the notation presented previously in the above discussion, especially in Section 3, is used in all cases, obtaining in this way a coherent description of every integral form usually needed in quantum chemical problems.

From now on it will be supposed that all functions appearing in the integrals here have a WO-CETO structure, defined in the sense of the discussion of Section 4, except for monocentric integrals, where the distinction between CETO and WO-CETO functions is irrelevant.

6.1. Overlap Integrals

a) One center.

The one center overlap integrals can be easily obtained considering equation (3.21):

$$S_{AA} = \langle\chi(\mathbf{A},\mathbf{a}',n',\alpha') \mid \chi(\mathbf{A},\mathbf{a}'',n'',\alpha'')\rangle = \langle\chi(\mathbf{A},\mathbf{a},n,\alpha)\rangle, \tag{6.1}$$

where $\mathbf{a}=\mathbf{a}'+\mathbf{a}''$, $n=n'+n''$, $\alpha=\alpha'+\alpha''$ and:

$$\langle\chi(\mathbf{A},\mathbf{a},n,\alpha)\rangle = \int \chi(\mathbf{A},\mathbf{a},n,\alpha)\, dV$$

$$= \delta(\Lambda_i(a_i=\dot{2}))\, F_1(a_1+a_2+1,a_3)\, F_2(a_2,a_1)\, t!\, \alpha^{-(t+1)}, \tag{6.2}$$

with $t=2+n+\Sigma_i a_i$.

The integral (6.2) vanishes unless $a_i=\dot{2}$, $\forall i$; and the function F_k, coming from the integration of the CETO function angular part, will be defined in section 6.1.c below.

b) Two center.

Taking into account the rotations outlined in section 3.1., the two-center overlap integral can be expressed as:

$$S_{AB} = \langle \chi(\mathbf{A},\mathbf{a},n,\alpha) \mid \chi(\mathbf{B},\mathbf{b},m,\beta) \rangle$$

$$= \Sigma(\mathbf{i}=\mathbf{0},\mathbf{a})\ \Sigma(\mathbf{j}=\mathbf{0},\mathbf{i})\ \Sigma(\mathbf{i}'=\mathbf{0},\mathbf{b})\ \Sigma(\mathbf{j}'=\mathbf{0},\mathbf{i}')$$

$$\bullet\ M(\mathbf{i},\mathbf{j},\mathbf{a})\ M(\mathbf{i}',\mathbf{j}',\mathbf{b})\ f(\mathbf{i},\mathbf{j},\mathbf{a},\varphi_A,\theta_A)\ f(\mathbf{i}',\mathbf{j}',\mathbf{b},\varphi_B,\theta_B)$$

$$\bullet\ \langle W(A,\mathbf{g},n,\alpha) \mid W(B,\mathbf{g}',m,\beta) \rangle, \qquad (6.3)$$

where it is necessary to compute overlap integrals over WO-CETO functions, as:

$$S^{w}_{AB} = \langle W(A,\mathbf{a},n,\alpha) \mid W(B,\mathbf{b},m,\beta) \rangle$$

$$= \delta(a_1+b_1=\dot{2}\wedge a_2+b_2=\dot{2})\ (R/2)^{t+1}\ F_2(a_2+b_2,a_1+b_1)$$

$$\bullet\ \Sigma(\mathbf{r}=\mathbf{0},\mathbf{t})\ \Gamma(\mathbf{r},\mathbf{c})\ A_{r_1}(\sigma)\ B_{r_2}(\tau), \qquad (6.4)$$

with $\mathbf{t}=(t,t)$, $t=2+m+n+\Sigma_i(a_i+b_i)$ and σ and τ are defined in equation (5.8). The coefficients $\Gamma(\mathbf{r},\mathbf{c})$ arise from the development of $P(\mathbf{c},u,v)$, with the definition of the six dimensional vector $\mathbf{c}=(n+1,m+1,a_3,b_3,k,k)$, as explained in section 3.4.3, see equations (3.36) to (3.39).

As it will be shown in next section c) below, the sum $a_1+b_1+a_2+b_2$ is even for the non-vanishing integrals, due to the properties of the function F_k, which is described there. Thus, the index

$k=(a_1+b_1+a_2+b_2)/2$, appearing in the expression of the vector **c** has always an integer value.

c) Auxiliary Overlap function [50].

Definition 9: The function $F_k(m,n)$.

In equations (6.2) and (6.4) the function $F_k(m,n)$ is defined as the integral:

$$F_k(m,n) = \int_0^{k\pi} (\sin x)^m (\cos x)^n \, dx$$

$$= \delta(n=\dot{2}) \{ k\delta(m=\dot{2}) (\Gamma(p)\Gamma(q)/\Gamma(p+q))$$

$$+ \delta(k=1 \wedge m\neq\dot{2}) (2^{m+1} m! \prod_{r=0}^{m} [2(n+r)+1]) \}, \qquad (6.5)$$

with $k\varepsilon\{1,2\}$, $p=(m+1)/2$, $q=(n+1)/2$ and $\Gamma(x)$ is the Gamma function.■

6.2. Kinetic Energy Integrals

a) Kinetic energy integrals may be computed in a similar way as the same kind of integrals over GTO's are, using overlap integrals and the symmetric form of the associated kinetic energy operator derived from Green's first identity [50b]:

$$T_{AB} = \langle \chi_A | -\tfrac{1}{2}\nabla^2 | \chi_B \rangle = \tfrac{1}{2} \langle \nabla \chi_A | \nabla \chi_B \rangle. \qquad (6.6)$$

Furthermore, using WO-CETO Definition 5, and applying the partial derivatives with respect the $\{x_i\}$ coordinates, it is obtained :

$$\begin{aligned} \partial W(A,\mathbf{a},n,\alpha) / \partial x_i = a_i\ & W(A,\mathbf{a}-\mathbf{e}_i,n,\alpha) \\ + n\ & W(A,\mathbf{a}+\mathbf{e}_i,n-2,\alpha) \\ - \alpha\ & W(A,\mathbf{a}+\mathbf{e}_i,n-1,\alpha), \end{aligned} \tag{6.7}$$

where $\{\mathbf{e}_i\}$ are the rows of the (3×3) unit matrix.

It must be noted here the fact that when a_i or n are zero in the original WO-CETO the affected terms in (6.7) vanish.

b) This expression can be used to describe not only kinetic energy integrals but as a starting point of a recurrent procedure to compute CETO integrals, in the same manner as Saunders did when dealing with GTO's [67]. Although this path will not be further explored here.

c) Then, the integral (6.6) can be expressed as a sum of three terms with respect of overlap integrals (6.4) as:

$$T_{AB} = \tfrac{1}{2} \Sigma_i\ T_{ABi}, \tag{6.8}$$

then employing the auxiliary definition:

$$K_i(s,p) = b_i \langle W(A,\mathbf{a}+s\mathbf{e}_i,n-p,\alpha) \mid W(B,\mathbf{b}-\mathbf{e}_i,m,\beta)\rangle + m \langle W(A,\mathbf{a}+s\mathbf{e}_i,n-p,\alpha) \mid W(B,\mathbf{b}+\mathbf{e}_i,m-2,\beta)\rangle - \beta \langle W(A,\mathbf{a}+s\mathbf{e}_i,n-p,\alpha) \mid W(B,\mathbf{b}+\mathbf{e}_i,m-1,\beta)\rangle, \quad (6.9)$$

where $s \in \{-1,+1\}$ is a sign and $p \in \{0,1,2\}$ an integer, one obtains:

$$T_{ABi} = a_i K_i(-1,0) + n K_i(+1,2) - \alpha K_i(+1,1). \quad (6.10)$$

d) When dealing with one center kinetic energy integrals, the term (6.10) can be written in the simplified form, using the WO-CETO combination:

$$\begin{aligned} U_{AAi} = & - a_i(a_i-1) && \chi(\mathbf{A},\mathbf{a}-2\mathbf{e}_i,n,\alpha) \\ & - n(2a_i+1) && \chi(\mathbf{A},\mathbf{a},n-2,\alpha) \\ & + \alpha(2a_i+1) && \chi(\mathbf{A},\mathbf{a},n-1,\alpha) \\ & - \delta(n>2)\, n(n-2) && \chi(\mathbf{A},\mathbf{a}+2\mathbf{e}_i,n-4,\alpha) \\ & + \alpha(n+(n-1)\delta(n>1)) && \chi(\mathbf{A},\mathbf{a}+2\mathbf{e}_i,n-3,\alpha) \\ & - \alpha^2 && \chi(\mathbf{A},\mathbf{a}+2\mathbf{e}_i,n-2,\alpha), \end{aligned} \quad (6.11)$$

then one can write:

$$T_{AAi} = \langle U_{AAi} \rangle, \quad (6.12)$$

where the symbol $\langle U_{AAi} \rangle$ stands for an integral like the one center overlap defined in equations (6.1,6.2) encompassing all the WO-CETO functions included in equation (6.11).

The same remarks on vanishing terms apply here in the same manner as when equation (6.7) is commented.

6.3. Variation of Mass with Velocity

The relativistic correction of the mass variation with velocity depends essentially on the fourth power of the nabla operator [68b]. In fact one can write the involved integral as:

$$R_{AB} = \langle \chi_A | \nabla^4 | \chi_B \rangle = \langle \nabla^2 \chi_A | \nabla^2 \chi_B \rangle, \tag{6.13}$$

Thus, the second derivatives of WO-CETO functions are needed. Using equation (6.7) where first derivatives are obtained, the second derivatives can be found by means of the six term WO-CETO combination shown in equation (6.11), that is:

$$\partial^2 W(A,\mathbf{a},n,\alpha) / \partial x_i^2 = -U_{AAi}. \tag{6.14}$$

As in equation (6.0) to (6.11) the integral R_{AB} can be written as a sum:

$$R_{AB} = \Sigma_i R_{ABi}, \tag{6.15}$$

where the three terms in the sum (6.15) will be obtained using the auxiliary expression involving six WO-CETO functions.

Using equation (6.14) each term in the sum (6.15) is obtained as an overlap integral:

$$R_{ABi} = \langle U_{AAi} \mid U_{BBi} \rangle, \tag{6.16}$$

As in the kinetic energy case, when one center mass variation integrals are to be evaluated, equation (6.16) can be written in a simplified form as occurred in equations (6.10) and (6.11).

6.4. Electric Moment Integrals

In order to compute these kind of integrals, let us allow any electric moment operator, belonging to the c-th order operator set, be referred to the center **C** and be written as:

$$\mathbf{M}_c(\mathbf{C},\mathbf{c}) = \prod_i (x_i - C_i)^{c_i}, \tag{6.17}$$

where $c=\Sigma_i c_i$.

The operator in equation (6.17) is equivalent to an angular CETO part expressible as $H_c(\mathbf{C},\mathbf{c})$.

After expanding the required binomial expansions, the electric moment integral can be written as:

$$\mathbf{M}_{AB} = \langle W(A,\mathbf{a},n,\alpha) \mid \mathbf{M}_c(C,\mathbf{c}) \mid W(B,\mathbf{b},m,\beta) \rangle$$

$$= \Sigma(\mathbf{k}=\mathbf{0},\mathbf{c})\, \Gamma(\mathbf{D},\mathbf{c},\mathbf{k}) \langle W(A,\mathbf{a}+\mathbf{k},n,\alpha) \mid W(B,\mathbf{b},m,\beta) \rangle, \qquad (6.18)$$

where $\Sigma(\mathbf{k}=\mathbf{0},\mathbf{c})$ is a nested summation symbol, and the coefficients $\Gamma(\mathbf{D},\mathbf{c},\mathbf{k})$ are defined as:

$$\Gamma(\mathbf{D},\mathbf{c},\mathbf{k}) = \prod_i (-1)^{c_i-k_i} \binom{c_i}{k_i} D_i^{c_i-k_i}, \qquad (6.19)$$

with **D=C-A**.

Thus, the CETO electric moment integrals are simple combinations of overlap integrals. A similar behavior as the one encountered in the GTO framework [68b] or in the STO case [18].

6.5. Some Remarks on One-electron Integrals

Some questions arise when considering the simple analytical forms of the integrals studied so far:

a) When dealing with two center kinetic energy integrals, overlap terms like $\langle W(A,\mathbf{a},p,\alpha) \mid W(B,\mathbf{b},q,\beta) \rangle$ can appear, where integer indices p or q can have values down to -1, as an effect of the derivation process. Computation of these integrals can be accomplished using equation (6.4). Negative integer exponent value -1 is not a problem upon integration, due to the fact that the differential volume in prolate

spherical coordinates has a term in $r_A \cdot r_B$, which cancels the previous effect. This property is also used when deriving two center nuclear attraction and Coulomb integral formulae.

b) Something similar occurs in the one center case. When observing integral terms like $\langle W(A,\mathbf{a},p,\alpha)\rangle$, the integer index p can take values as low as : -1 or -2, but the volume differential in spherical coordinates has a term in r^2, which cancels the previous negative power of r.

c) The appearance of higher negative powers of r will be a dominant problem when dealing with higher derivatives of WO-CETO functions as in the variation of mass with velocity integrals and Coulomb-like operators. Next section 7 takes care of some aspects facing this problem.

6.6. Nuclear Attraction Integrals

Nuclear attraction integrals being the first touchstones of the present procedure, are discussed here, at this section end, opening the way to two electron integrals problem which will be discussed next in section 7.

There are four kinds of nuclear attraction integrals. Three of them do not need the ε_{AB} expansion (5.2) and are presented first.

Because of the many center nature of the fourth integral case, a detailed analysis of three center nuclear attraction integral problem is given. Using the ideas developed in Sections 4 and 5, it is described how the three center integrals become expressible in terms of one and two center ones. An example involving s-type WO-CETO functions is presented as a test of the developed theory of the preceding chapters.

6.6.1. One Center Integrals of the form <AA:A>

These type of integrals can be computed integrating within a spherical coordinates framework. They can be associated to a one center overlap integral (6.2) as follows:

$$\begin{aligned} \langle AA':A\rangle &= \int \chi(\mathbf{A},\mathbf{a},n,\alpha)\, r^{-1}\, \chi(\mathbf{A},\mathbf{a}',n',\alpha')\, dV \\ &= \langle \chi(\mathbf{A},\mathbf{a}+\mathbf{a}',n+n'-1,\alpha+\alpha')\rangle. \end{aligned} \tag{6.20}$$

Again a well known result is met, resembling the usual pattern found in the STO or GTO choices.

6.6.2. Two Center Integrals of the form <AA:B>.

These type of integrals and the following cases, can be obtained integrating in a prolate spherical coordinates framework.

a) When dealing with two or more centers, some rotations are needed in order to perform the integration upon WO-CETO functions. Thus, in this case, one can see that the nuclear attraction integral is a finite linear combination of two center overlap ones:

$$\begin{aligned}
\langle AA':B\rangle &= \int \chi(\mathbf{A},\mathbf{a}',n',\alpha')\, r_B^{-1}\, \chi(\mathbf{A},\mathbf{a}'',n'',\alpha'')\, dV \\
&= \int \chi(\mathbf{A},\mathbf{a},n,\alpha)\, r_B^{-1}\, dV = \langle \chi(\mathbf{A},\mathbf{a},n,\alpha) \mid r_B^{-1} \rangle \\
&= \Sigma(\mathbf{i}=\mathbf{0},\mathbf{a})\, \Sigma(\mathbf{j}=\mathbf{0},\mathbf{i})\, M(\mathbf{i},\mathbf{j},\mathbf{a})\, f(\mathbf{i},\mathbf{j},\mathbf{a},\varphi,\theta) \\
&\bullet \langle W(A,\mathbf{g},n,\alpha) \mid W(B,\mathbf{0},-1,0) \rangle, \qquad (6.21)
\end{aligned}$$

where $\mathbf{a}=\mathbf{a}'+\mathbf{a}''$, $n=n'+n''$ and $\alpha=\alpha'+\alpha''$. The M and f coefficients and the vectors $\mathbf{g}=\mathbf{g}(\mathbf{a},\mathbf{i},\mathbf{j})$ constructed as equation (3.9), arise from the rotations transforming the CETO centered at **A** into a well-oriented form with respect the CETO centered at **B**.

Also the term $r_B^{-1}=\chi(\mathbf{B},\mathbf{0},-1,0)=W(B,\mathbf{0},-1,0)$ can be expressed either as a CETO or a WO-CETO, relaxing condition 2) and 3) of CETO Definition 1. In this manner the last CETO representing the operator r_B^{-1} does not have to be rotated because it only has a radial part.

b) The overlap-like integrals appearing in the equation (6.21) which involve the operator W(B,**0**,-1,0), can be taken as a particular case of the two center overlap integrals described in equation (6.4). They have the following form:

$$\langle W(A,\mathbf{a},n,\alpha) \mid W(B,\mathbf{0},-1,0)\rangle = (R/2)^{t+1}\, F_2(a_2,a_1)$$

$$\cdot\, \Sigma(\mathbf{r}=\mathbf{0},\mathbf{t})\, \Gamma(\mathbf{r},\mathbf{c})\, A_{r_1}(\sigma)\, B_{r_2}(\sigma), \quad (6.22)$$

with $\mathbf{t}=(t,t)$, $t=1+n+\Sigma_i a_i$ and $\sigma=\alpha R/2$. Again, the coefficients $\Gamma(\mathbf{r},\mathbf{c})$ arise from the development of $P(\mathbf{c},u,v)$, with the vector $\mathbf{c}=(1+n,0,a_3,0,k,k)$ and the index $k=(a_1+a_2)/2$, as it is explained in section 3.4.3.b.

c) From another point of view, the integral (6.22) can be considered as an infinite linear combination of s-type ETO's.

Because the integrals $B_k(x)$ may be computed as a power series of the argument x, as it is shown in equation (3.48), it may be also interesting to obtain the expression (6.22) as an infinite power series of the intercenter distance R, that is:

$$\langle W(A,\mathbf{a},n,\alpha) \mid W(B,\mathbf{0},-1,0)\rangle =$$

$$F_2(a_2,a_1)\, 2^{-t}\, e^{-\alpha R/2}\, \Sigma(\mathbf{i}=\mathbf{0},\mathbf{N})\, \Omega(\mathbf{i},\mathbf{c},\alpha)\, R^p, \quad (6.23)$$

where in equation (6.23) the same definitions as in equation (6.22) are used, with $\mathbf{N}=(\infty,t,t,i_2)$, $p=t+q$ where $q=2i_1-i_2+i_4+d$ and d is a logical Kronecker delta defined as $d=\delta(i_3\neq\dot{2})$.

The Ω coefficients are constructed in turn as:

$$\Omega(\mathbf{i},\mathbf{c},\alpha) = \Gamma((i_2,i_3),\mathbf{c})\, (\alpha/2)^{q-1}\, w(i_1,i_3,d)\, i_2!/i_4!, \quad (6.24)$$

and the w coefficients are obtained by means of:

$$w(i, j,d) = (-1)^d / \{ (2i+d)! \, (2(i+[\![j/2]\!]+d)+1) \}. \qquad (6.25)$$

The same can also be said of the overlap expression written in equation (6.4), due to the presence of the B integrals: An infinite power series similar to the equation (6.23) can be easily deduced.

6.6.3. Two Center Integrals of the form <AB:A> or <AB:B>

In this case, one has:

$$\begin{aligned} \langle AB:A \rangle &= \int \chi(\mathbf{A},\mathbf{a},n,\alpha)\, r_A^{-1}\, \chi(\mathbf{B},\mathbf{b},m,\beta)\, dV \\ &= \langle \chi(\mathbf{A},\mathbf{a},n-1,\alpha) \,|\, \chi(\mathbf{B},\mathbf{b},m,\beta) \rangle. \end{aligned} \qquad (6.26)$$

Thus, a finite linear combination of WO-CETO overlap integrals suffice to compute the two center <AB:A> one.

6.6.4. Three Center Integrals of the form <AB:C>

Now, one has to solve in this case the integral form:

$$<AB:C> = \int \chi(\mathbf{A},\mathbf{a},n,\alpha)\ r_C^{-1}\ \chi(\mathbf{B},\mathbf{b},m,\beta)\ dV$$

$$= <\chi(\mathbf{A},\mathbf{a},n,\alpha) \mid r_C^{-1} \mid \chi(\mathbf{B},\mathbf{b},m,\beta)>$$

$$= <\Omega_{AB} \mid r_C^{-1}>. \qquad (6.27)$$

Here, the expansion expressed in equation (4.4) with the ε_{AB} functions defined as in equation (5.1) are needed in order to perform the computation of the integral (6.27), and one can write:

$$<AB:C> = \sum_{\alpha\in A}\ c_\alpha^A\ <\psi_\alpha^A \mid r_C^{-1}> + \sum_{\beta\in B}\ c_\beta^B\ <\psi_\beta^B \mid r_C^{-1}>. \qquad (6.28)$$

One can take into account that in expression (6.28) above the functions $\{\psi_\alpha^A,\psi_\beta^B\}$ are WO-CETO functions referred to the centers **A** and **B** respectively, but behave as CETO functions when each one in turn is associated with the operator $\mid r_C^{-1}>$, which may be also considered as having a WO-CETO function form W(C,**0**,-1,0). Then, if the symbols $<\psi_\delta^D \mid r_C^{-1}>$ are associated with the corresponding integrals and $D\in\{A,B\}$, $\delta\in\{\alpha,\beta\}$, the following equivalence can be written:

$$<\psi_\delta^D \mid r_C^{-1}> = <\chi(\mathbf{D},\mathbf{d}_\delta,n_\delta,\tau_\delta) \mid r_C^{-1}>. \qquad (6.29)$$

One encounters an expression similar to the one appearing in the Nuclear Attraction Integrals of the same kind as these discussed in section 6.2, as shown in equation (6.21).

6.6.5. Some Computational Results on Nuclear Attraction Integrals

Some results obtained applying equation (6.28) are presented in **Table 6.1.**

Table 6.1

Variation of the Three Center Nuclear Attraction Integrals with L_n(*).

Nuclear Configurations and Values given by Steinborn [37a] with $\alpha=\beta=1$:
X : C -- 0.5 -- A -- 2.0 -- B = 0.341022
Y : A -- 0.5 -- C -- 1.5 -- B = 0.449956

Number of Functions n	X Integral	Y Integral
1	0.420495	0.478926
2	0.352532	0.463547
3	0.349553	0.459502
4	0.349028	0.458835
5	0.339212	0.448500
6	0.338475	0.447857
7	0.353745	0.463101
8	0.341179	0.450433
9	0.341341	0.450259
10	0.341302	0.450215
11	0.341277	0.450187
12	0.340975	0.449612
13	0.341048	0.449980
14	0.341045	0.449978
15	0.341023	0.449958

(*) using ε_{AB} expansion optimized functions of **Table 5.1**

It can be seen that on the contrary of Steinborn's [37a] computation, involving a linear combination of more than 150

functions, of the same integral values, a small number of CETO's gives the desired result within a reasonable accuracy.

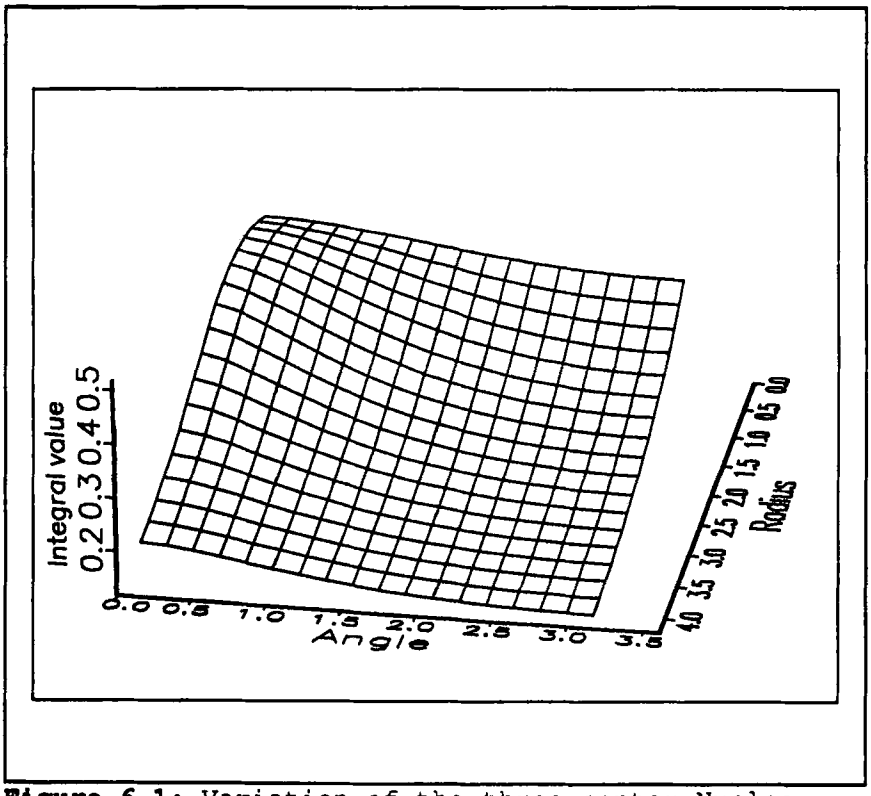


Figure 6.1: Variation of the three center Nuclear Attraction integral $\langle 1s_A 1s_B : C \rangle$ value.

Figure 6.1 shows the three center nuclear attraction integral variation using the same ε_{AB} covering of **Figure 5.1**, and the same

parameters. Center **C** is moved around the **A-B** fixed nuclear position and varied with respect a radius vector measured from **A** and an angle **CAB** in radians. The expected integral behavior is reflected in the surface shape.

7. Integrals Over Two-electron Operators

Repulsion integrals, being the real stars of the present procedure, conserving the important role already obtained in GTO and STO choices, are discussed here, in this closing section.

The two electron integral problem is reviewed in a general context, in order to elucidate the integral forms needed when using CETO products. The result, already advanced schematically in section 4.3 indicates that in any case, when employing the techniques previously discussed, only one and two center bielectronic integrals are to be used.

From this results as starting point, the one and two center Coulomb integral computation is set, using the structure of the nuclear attraction monoelectronic integrals to solve the integration over the coordinates of the first electron, as has been done early in classical work [47], or reversing this integration scheme as O-ohata and Ruedenberg did, within the STO functions framework, in an interesting paper published in 1966 [26b].

The final labor, the integral over the coordinates of the second electron, in this CETO function case, corresponds to a calculation of integrals, involving in the worse situation an infinite series, by means of the practical use of the ideas discussed in section 3.4.

This section also includes at the end an analysis of the electron spin-spin contact integrals, a well known and simple relativistic correction term, which are also studied in this paper due to their close relationship with Quantum Similarity Measures [66c]. The form of such integrals corresponds to some kind of many function, many center overlap.

Other relativistic correction integrals are also discussed as a general extension of one and two electron Coulomb operators, in a context where it is discussed a scheme to deal with one and two electron integrals using a common point of view.

7.1. General Considerations on Two-electron Operators and Integrals

a) Let us write a large set of integrals involving two density functions and a two electron operator $T(\mathbf{r}_1,\mathbf{r}_2)=\mathbf{T}_{12}$ as:

$$\langle AB \mid \mathbf{T} \mid CD \rangle = \iint \Omega_{AB}(1)\, \mathbf{T}_{12}\, \Omega_{CD}(2)\ dV_1 dV_2. \qquad (7.1)$$

Thus, four CETO's are involved which, when discussing one electron operator integrals as before, can be considered as a linear combination of WO-CETO's when A≠B or C≠D and can be left as CETO functions when studying the integral (7.1) in the one center case.

b) Repulsion as well as spin-spin contact integrals will be discussed next. In both cases, in fact in any case when dealing with integrals like (7.1) and thinking of the main procedure where CETO integral calculation can rely, it is necessary to take into account the four situations where A=B or A≠B and C=B or C≠B. That is because, after the discussion in section 4 above, originated by equation (4.4) or (4.5) a two electron integral like (7.1) can be written as:

$$\begin{aligned} \langle AB \mid \mathbf{T} \mid CD \rangle \approx & \sum_a \sum_c C_a^{AC} C_c^{CA} \langle a \mid \mathbf{T} \mid c \rangle \\ & + \sum_a \sum_d C_a^{AD} C_d^{DA} \langle a \mid \mathbf{T} \mid d \rangle \\ & + \sum_b \sum_c C_b^{BC} C_c^{CB} \langle b \mid \mathbf{T} \mid c \rangle \\ & + \sum_b \sum_d C_b^{BD} C_d^{DB} \langle b \mid \mathbf{T} \mid d \rangle . \end{aligned} \tag{7.2}$$

c) An interesting case, which can be easily solved within the present approach corresponds to the two center integrals of "exchange" type: $\langle AB \mid \mathbf{T} \mid AB \rangle$. It is well known that in the STO framework and with the Coulomb operator substituting the operator **T**, they are not so trivially computed as the two center "Coulomb" type integrals: $\langle AA \mid \mathbf{T} \mid BB \rangle$, as Ruedenberg's [14a] or Wahl et al. [47] papers prove. But using equation (7.2) they are simply expressed in terms of one center: $\langle A \mid \mathbf{T} \mid A' \rangle$, $\langle B \mid \mathbf{T} \mid B' \rangle$ and two center: $\langle A \mid \mathbf{T} \mid B \rangle$ integrals. The same can be said for the "hybrid" type integrals $\langle AA \mid \mathbf{T} \mid AB \rangle$.

7.1.1. Reduced Two-electron Integrals

a) Thus, only one and two center integrals are to be computed mainly because $\Omega_{AA'}$ densities can be also considered, according equation (3.21), as a new CETO centered at A. Let us call the $\langle A|\mathbf{T}|B\rangle$ integrals **reduced two-electron integrals**. They are nothing but integrals which can be expressed according:

Definition 10: Reduced two-electron integrals.

$$\langle A|\mathbf{T}|B\rangle = \iint \chi_A(1)\, \mathbf{T}_{12}\, \chi_B(2)\, dV_1 dV_2, \tag{7.3}$$

where $\chi_A(1)$ and $\chi_B(2)$ are ETO functions centered on sites **A** and **B** respectively and $\mathbf{T}_{12}$ corresponds to any two electron operator.■

In this sense, electronic spin-spin contact integrals can be computed as simple bilinear functions of overlap integrals, and electron repulsion integrals are expressible as bilinear functions of some sort of two electron nuclear attraction integrals.

b) The problem of CETO orientation will appear again in this kind of integrals, but each reduced integral can be supposed composed of WO-CETO's which can be further transformed into CETO functions by an appropriate rotation scheme.

One must consider that functions {| A>, | B>} are WO-CETO's with respect the AB axis, and the same but for the CD axis when functions {| C>, | D>} are studied. Hence reduced two-electron integrals of any kind: <A | T | B>, ... in equation (7.2) are to be considered over WO-CETO's, which in turn can also be taken as E-CETO's, needing to be transformed afterwards by applying the procedure outlined in section 3.1.c. Thus, only reduced integrals over WO-CETO functions have to be computed.

7.2. Repulsion Integrals

In this case the two electron operator is: $T_{12}=r_{12}^{-1}$ and according to the previous general discussion on two electron integrals, only two kinds of **reduced repulsion integrals** over WO-CETO functions should be taken into account: one and two center integrals.

7.2.1. Integration Framework

The general integration structure adopted here in order to obtain analytic formulae is based on the first integration associated to a nuclear attraction integral like these appearing in equation (6.21). A similar strategy, which seems the most natural way to proceed, was also used by Wahl et al. [47]. In fact, writing the following general **reduced Coulomb integral**:

$$R_{AB} = \langle A | r_{12}^{-1} | B \rangle$$
$$= \iint [W(A,\mathbf{a},n,\alpha)\{1\}]\, r_{12}^{-1} [W(B,\mathbf{b},m,\beta)\{2\}]\, dV_1 dV_2. \qquad (7.4)$$

It is straightforward to see how integration with respect to the electron {2}, for example, is equivalent to compute a nuclear attraction integral like:

$$\vartheta(1) = [\langle B | r_{12}^{-1} \rangle \{1\}] = [\langle W(B,\mathbf{b},m,\beta) | r_{12}^{-1} \rangle \{1\}], \qquad (7.5)$$

which in turn is a function of electron {1} coordinates. Thus, the reduced repulsion integral may be written:

$$R_{AB} = \int [W(A,\mathbf{a},n,\alpha)\{1\}]\, \vartheta(1)\, dV_1. \qquad (7.6)$$

Then, reduced repulsion integrals have similar forms in both one and two center case.

The same analysis can be applied when considering CETO functions instead of WO-CETO's in the integral (7.4) and the following equations (7.5) and (7.6). Moreover, the same occurs when dealing with densities in general or STO densities in particular.

The integration in equation (7.6), over the coordinates of the remaining electron {1} was performed numerically by Wahl et al. in

reference [47]. The reason forcing a numerical integration procedure at the second stage may be possibly due to the fact that integration over the potential, even constructed from CETO functions using the preceding discussion and the next section technique, has an analytical form which in some cases produce a set of divergent integrals. This drawback will be discussed below.

7.2.2. Potential Functions

The integration over the coordinates of the electron number (2) will be done over a prolate spherical coordinate system $\{r_{B2}, r_{12}, r_{B1}\}$, and the analytic expression of the nuclear attraction integral shown in equation (7.5) obtained. This integration method ensures that the variable r_{12} will not produce any mathematical problem as it has been observed in section 6.5. Note that the constant parameter, which is the usual intercenter distance, when dealing with this kind of coordinate system, will act here as a variable and it has to be integrated when considering the electron (1) coordinates.

In order to have electron (2) in a well-oriented system with respect to electron (1), it is necessary to perform a rotation, for each position of electron (1), from the original system to the prolate spherical system described in the paragraph 7.2.1.

Let the WO-CETO attached to electron (2) coordinates be symbolized by: $W(B,\mathbf{b},m,\beta)$.

Electrons (1) and (2) are described by means of the spherical coordinate systems $\{r_{A1},\theta_1,\varphi_1\}$ and $\{r_{B2},\theta_2,\varphi_2\}$, respectively.

After performing the needed rotations outlined in section 3.1.b, one can express the $\vartheta(1)$ potential function (7.5) as:

$$\vartheta(1) = \Sigma(\mathbf{i=0,b})\ \Sigma(\mathbf{j=0,i})\ M(\mathbf{i,\ j,\ b})\ f(\mathbf{i,\ j,\ b},\varphi_1,\theta_{B1}) \cdot \langle W(B,\mathbf{g},m,\beta)\,|\,W(\{1\},\mathbf{0},-1,0)\rangle, \tag{7.7}$$

where $\Sigma(\mathbf{i=0,b})$ and $\Sigma(\mathbf{j=0,i})$ are nested summation symbols, defined as in section 3.1 as well as the coefficients $M(\mathbf{i,\ j,\ b})$ and using the function $f(\mathbf{i,\ j,\ b},\varphi_1,\theta_{B1})$ which appears in equation (3.5) for the first time. Note that the angle θ_{B1} is measured from the center **B** and it refers to the electron (1).

The vector $\mathbf{g=g(b,\ i,\ j)}$ has the same definition as in equation (3.9).

Using equation (6.21) and the expression (3.49) related to the product of A and B integrals, equation (7.7) uses:

$$\langle W(B,\mathbf{g},m,\beta)\,|\,W(\{1\},\mathbf{0},-1,0)\rangle = F_2(g_2,g_1)\ N(T_b,\beta) \cdot \Sigma(\mathbf{p=0,q})\ A(\mathbf{p,h},\beta)\ r_{B1}^{T_b-s}\ \{(-1)^{p_4}-\exp(-\beta r_{B1})\}. \tag{7.8}$$

Where:

a) $T_b=m+\Sigma_i b_i$.

b) The function N is defined by: $N(t,\beta) = 2^{-t}\ \beta^{-2}$. (7.9)

c) The coefficients A are constructed by means of:

$$\mathbf{A(p,h},\beta) = \Gamma(\mathbf{p,h})\ \Lambda(\mathbf{p})\ (2/\beta)^{s}. \tag{7.10}$$

where Γ appears as in equation (6.4), being described in equations (3.37) to (3.39), also the Λ coefficients are defined in equation (3.49) and $s=p_1+p_2-(p_3+p_4)$.

d) The six components of the vector **h** are:

$$\mathbf{h} = (1+m,0,g_3,0,k,k), \tag{7.11}$$

with $k=(g_1+g_2)/2$, and are obtained from the expansion of an expression similar to the one outlined in equation (3.35) which leads to a formula like (3.36).

e) Finally: $\mathbf{q} = (m+1,m+1,p_1,p_2)$.

Although the previous scheme seems very promising for both one and two center cases, we faced the same difficulties as Wahl et al. [47] surely have met, when integration over the second electron was to be

performed. In consequence, we use this approach only for the one center integrals, discussed next.

7.3. One Center Reduced Coulomb Integrals

Let W(1)=W(A,**a**,n,α) and W(2)=W(B,**b**,m,β) be two WO-CETO's centered at the origin, that is: **A=B=0**.

One can define the following quantities:

a) $T_a=n+\Sigma_i a_i$ and $T_b=m+\Sigma_i b_i$ as the total order of each WO-CETO.
b) $T=T_a+T_b$ as the total order of the integral.

The integral to be calculated is:

$$R_{AA} = \langle A | r_{12}^{-1} | A \rangle = \iint W(1)\, r_{12}^{-1} W(2)\, dV_1 dV_2, \tag{7.12}$$

which can be written using the potential integral (7.5) expressed here as:

$$\vartheta(1) = \int r_{12}^{-1} x_1(2)^{b_1} x_2(2)^{b_2} x_3(2)^{b_3} r_2^m \exp(-\beta r_2)\, dV_2. \tag{7.13}$$

Thus, the integral (7.12) is:

$$R_{AA} = \int x_1(1)^{a_1} x_2(1)^{a_2} x_3(1)^{a_3} r_1^n \exp(-\alpha r_1)\, \vartheta(1)\, dV_1. \tag{7.14}$$

One can obtain the expression of $\vartheta(1)$ using the equation (7.8) where the needed prolate spherical system, see Definition 7, is configured by means of the parameters $\{r_2,r_{12},r_1\}$.

Finally an integration over the electron {1} is carried out in a spherical coordinates framework $\{r_1,\theta_1,\varphi_1\}$.

The final result is:

$$R_{AA} = \delta(\Lambda_i(a_i+b_i=\dot{2}))\ N(T_b\text{-}1,\beta)\ \Sigma(\mathbf{i=0,b})\ \Sigma(\mathbf{j=0,i})$$
$$\bullet\ C(\mathbf{i,a,b})\ \Sigma(\mathbf{p=0,q})\ A(\mathbf{p,h},\beta)\ D(\mathbf{p},n,\alpha,\beta). \qquad (7.15)$$

Where:

a) The coefficients C are defined as:

$$C(\mathbf{i,a,b}) = B(\mathbf{i,a,b})\ F_1(a_1+a_2+i_1+i_2+i_3\text{-}(j_1+j_2)+1,a_3+b_3+j_1+j_2\text{-}i_3). \qquad (7.16)$$

with the coefficients B being defined in turn by means of:

$$B(\mathbf{i,a,b}) = M(\mathbf{i,\ j,\ b})\ F_2(b_1+b_2\text{-}(i_1+i_2),j_1+j_2+i_3)$$
$$\bullet\ F_2(a_2+b_1\text{-}i_1+i_2,a_1+b_2+i_1\text{-}i_2). \qquad (7.17)$$

where $M(\mathbf{i,\ j,\ b})$ are the usual transformation coefficients,

b) N and A are the same as the coefficients defined in the equations (7.9) and (7.10) respectively, and the functions F_k are described in equation (6.5), Definition 9.

c) The six components of the vector **h** are the same as these defined in the equation (7.11).

d) The vector **q** has their components constructed as: (T_b+1,T_b+1,p_1,p_2).

e) The coefficients D are defined as:

$$D(\mathbf{p},n,\alpha,\beta) = (t-1)!\ \{(-1)^{p_4}\ \alpha^{-t} - (\alpha+\beta)^{-t}\}, \qquad (7.18)$$

with $t=n-(p_1+p_2)+(p_3+p_4)+3$.

f) From the parity restrictions associated to the F_k function in Definition 9, it is easy to determine that the following conditions hold necessarily:

$$\{a_1+b_1=\dot{2}\ \wedge\ a_2+b_2=\dot{2}\ \wedge\ a_3+b_3=\dot{2}\}, \qquad (7.19)$$

in order to have non-vanishing integrals. This condition generates the logical Kronecker delta appearing in equation (7.15). A similar situation is also found when dealing with GTO's [69].

g) Before carrying out any integral calculation of the kind described in this section, it is also necessary to ensure that $T_a \geq T_b$. That is because when obtaining $\vartheta(1)$, after an application of equation (7.8), some negative powers of r_1 appear, being the maximum absolute value $1+\Sigma_i b_i$. So, this condition is necessary in order to annul these negative powers when introducing $\vartheta(1)$ into equation (7.14) being as a result r_1 powered to positive numbers only, and doing so the calculation is simplified.

7.4. Two Center Reduced Coulomb Integrals

Let W_A, W_B be two WO-CETO functions placed in the centers **A** and **B** respectively, with **A≠B**, and being R the intercenter distance.

Due to the problems discussed at the end of section 7.2.2., the integration in the two center case has been done using the methodology outlined by O-Ohata and Ruedenberg [26b]. The integral to be computed now is a **reduced Coulomb integral** written as:

$$R_{AB} = \iint W_A(\mathbf{r}_1-\mathbf{r}_A)\ |\mathbf{r}_1-\mathbf{r}_2|^{-1}\ W_B(\mathbf{r}_2-\mathbf{r}_B)\ d\mathbf{r}_1 d\mathbf{r}_2, \qquad (7.20)$$

and making the change of variable, according to O-ohata and Ruedenberg [26b]:

$$\mathbf{r}_2-\mathbf{r}_B = \mathbf{r}_1-\mathbf{r}_C, \qquad (7.21)$$

one has to evaluate the integral (7.21) as:

$$R_{AB} = \iint W_A(\mathbf{r}_1-\mathbf{r}_A)\ |\mathbf{r}_C-\mathbf{r}_B|^{-1}\ W_B(\mathbf{r}_1-\mathbf{r}_C)\ d\mathbf{r}_1 d\mathbf{r}_C, \tag{7.22}$$

or

$$R_{AB} = \int |\mathbf{r}_C-\mathbf{r}_B|^{-1}\ S(\mathbf{r}_C-\mathbf{r}_A)\ d\mathbf{r}_C, \tag{7.23}$$

where an overlap integral has to be calculated between the CETO functions χ_A, χ_C, resulting from the translation of W_A, W_B to the new frame defined in equation (7.21):

$$S(\mathbf{r}_C-\mathbf{r}_A) = S(\mathbf{r}_{AC}) = \int \chi_A(\mathbf{r}_1-\mathbf{r}_A)\ \chi_C(\mathbf{r}_1-\mathbf{r}_C)\ d\mathbf{r}_1$$

$$= \Sigma(\mathbf{i}=\mathbf{0},\mathbf{a})\ \Sigma(\mathbf{j}=\mathbf{0},\mathbf{i})\ \Sigma(\mathbf{i}'=\mathbf{0},\mathbf{b})\ \Sigma(\mathbf{j}'=\mathbf{0},\mathbf{i}')$$

$$\bullet\ M(\mathbf{i},\mathbf{j},\mathbf{a})\ M(\mathbf{i}',\mathbf{j}',\mathbf{b})$$

$$\bullet\ f(\mathbf{i},\mathbf{j},\mathbf{a},\varphi_A,\theta_A)\ f(\mathbf{i}',\mathbf{j}',\mathbf{b},\varphi_B,\theta_B)$$

$$\bullet\ \langle W(A,\mathbf{g},n,\alpha)\,|\,W(C,\mathbf{g}',m,\beta)\rangle, \tag{7.24}$$

being $\mathbf{g}=\mathbf{g}(\mathbf{a},\mathbf{i},\mathbf{j})$ and $\mathbf{g}'=\mathbf{g}(\mathbf{b},\mathbf{i}',\mathbf{j}')$ as they are defined in equation (3.9).

Using the expression (7.24) in equation (7.23) and the Laplace expansion of $|\mathbf{r}_C-\mathbf{r}_B|^{-1}$, one arrives to the final result:

$$R_{AB} = \delta(a_1+b_1=\dot{2} \wedge a_2+b_2=\dot{2})$$

$$\bullet\ R^2\,(R/2)^{\iota+1}\,\Sigma(\mathbf{i}=\mathbf{0},\mathbf{a})\,\Sigma(\mathbf{j}=\mathbf{0},\mathbf{i})\,\Sigma(\mathbf{i}'=\mathbf{0},\mathbf{b})\,\Sigma(\mathbf{j}'=\mathbf{0},\mathbf{i}')$$

$$\bullet\ \mathbf{M(i,j,a)}\,\mathbf{M(i',j',b)}\,F_2(g_2+g_2',g_1+g_1')\,F_2(c_1+c_1',c_2+c_2')$$

$$\bullet\ \Sigma(\mathbf{r}=\mathbf{0},\mathbf{t})\,\Gamma(\mathbf{r},\mathbf{d})\sum_{s=0}^{\infty} P(s,c_3+c_3',c_4+c_4')\,J(\mathbf{u},\sigma,\tau), \qquad (7.25)$$

where:

a) The **c** and **c'** vectors are defined by means of the general vector structure:

$$\mathbf{c} = \mathbf{c(a,i,j)} = (a_1-i_1+i_2, a_2+i_1-i_2, i_1+i_2+i_3-(j_1+j_2), a_3+j_1+j_2-i_3), \qquad (7.26)$$

being $\mathbf{c}'=\mathbf{c(b,i',j')}$ defined in a similar manner.

b) The function P is:

$$P(\mathbf{k},m,n) = \int_0^{\pi} P_k(\cos\theta)\,(\sin\theta)^m\,(\cos\theta)^n\,d\theta$$

$$= 2^{-k}\sum_{p=0}^{q}(-1)^p\,(2(k-p))!/(p!\,(k-p)!\,(k-2p)!)\,F_1(m,k+n-2p), \qquad (7.27)$$

where $q=[\![k/2]\!]$ and being $P_k(x)$ Legendre polynomials of order k [50].

c) $$J(\mathbf{u},\sigma,\tau) = K(\mathbf{u},\sigma,\tau) + L(\mathbf{u},\sigma,\tau), \qquad (7.28)$$

where $\mathbf{u}=(\mathbf{r},s,t)$, $\mathbf{r}=(r_1,r_2)$ and:

$$K(\mathbf{u},\sigma,\tau) = \int_1^{\infty} x^{(t-s+2)}\, A_{r_1}(\sigma x)\, B_{r_2}(\tau x)\, dx,$$
$$L(\mathbf{u},\sigma,\tau) = \int_0^{1} x^{(t+s+3)}\, A_{r_1}(\sigma x)\, B_{r_2}(\tau x)\, dx. \qquad (7.29)$$

One can use the expression (3.52) to develop the product $A_k(x)B_l(y)$ in equations (7.29) avoiding negative powers of x in the function $L(\mathbf{u},\sigma,\tau)$. In this manner, the K integrals lead to compute integrals like $C_k(t)$ as defined in equation (3.53) and the L integrals need the computation of D_k integrals as defined in equation (3.56).

In this formulation $\mathbf{t}=(t,t)$, $t=2+m+n+\Sigma(a_i+b_i)$, the vector $\mathbf{d}=(n+1,m+1,g_3,g_3',h,h)$ is constructed with $h=(g_1+g_1'+g_2+g_2')/2$ and σ and τ are defined in the equation (5.8).

From the parity restrictions associated to the F_k function of Definition 9, it is easy to determine that the following conditions must hold simultaneously:

$$\{a_1+b_1=\dot{2} \wedge a_2+b_2=\dot{2}\} \qquad (7.30)$$

in order to have non-vanishing integrals and this is the origin of the logical Kronecker delta appearing in equation (7.25).

7.5. Electron Spin-Spin Contact Integrals

In this case the two-electron operator is defined by means of a three dimensional Dirac delta function:

$$\mathbf{T}_{12} = \delta(\mathbf{r}_1-\mathbf{r}_2) = \prod_i \delta(x_i(1)-x_i(2)), \tag{7.31}$$

and thus:

$$\langle AB \mid \mathbf{T} \mid CD\rangle = \int \Omega_{AB}(\mathbf{r}_1)\, \delta(\mathbf{r}_1-\mathbf{r}_2)\, \Omega_{CD}(\mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2$$

$$= \int \Omega_{AB}(\mathbf{r})\, \Omega_{CD}(\mathbf{r})\, d\mathbf{r} = \langle AB:CD\rangle. \tag{7.32}$$

Then, in general, all integrals of this kind become functions of one electron overlap integrals. In the most difficult case, where A≠B≠C≠D, the adaptation to various well oriented CETO pairs is inevitable but as the integral is one-electron in practice one can choose any of the AB, AC, AD, ... pairs or whichever conveniently related form to develop the integral transformation. That is because the 24 permutations of the four involved CETO functions leave the integral invariant, that is:

$$\langle AB:CD\rangle = \langle DA:BC\rangle = \langle CD:AB\rangle = \langle BC:DA\rangle = \langle DB:CA\rangle = \ldots$$

Thus, the expansion (4.4) as well as overlap evaluation and transformation rules are sufficient to compute electron spin-spin contact integrals.

These results permit one to see that the same will occur when the integral of an arbitrary number of CETO products centered at arbitrary sites $\{A_i\}$ is to be evaluated. Any permutation of the functions in the product will leave the integral invariant and WO-CETO transformation rules as well as overlap integral evaluation will be sufficient to obtain this kind of integrals. These characteristics, which are also present when using GTO functions, may appear interesting when dealing with Quantum Similarity measures [66b,d].

7.6. Calculations

Some examples of computed bielectronic repulsion integral values are presented in tables **7.1** and **7.2**.

Table 7.1 shows how repulsion integrals over CETO functions can be constructed with reliable accuracy, as three center nuclear attraction integrals were computed.

Table 7.2 shows a simple set of repulsion integrals. All the involved functions appearing in **Table 7.2** are 1s functions with unit scaling factors.

Table 7.1

Integral type	Number of functions in L_n: 5	10	15	20	Ref. value	Ref.
(AA,AB)[a]	0.18377	0.18389	0.17971	0.16092	0.17000	[70]
(AA,AB)[b]	0.50654	0.50339	0.50703	0.50702	0.50704	[71]
(AB,AB)[c]	0.43578	0.43019	0.43663	0.43655	0.43665	[71]
(AB,AB)[d]	0.01825	0.01801	0.01800	0.01800	0.01800	[70]
(AB,BC)[e]	0.16260	0.16520	0.16587	0.16582	0.16985	[71]

a) $\times 10^3$. From Hurley [70]: Integral appearing in a carbon monoxide calculation. R_{CO}=2.1319 au. A=$1s_C(\alpha=5.67)$, B=$1s_O(\alpha=7.66)$.

b) and c) From Hirschfelder [71]: Integrals required for computing the energy of H_3 and of H_3^+. R_{AB}=1.0 au, all the CETO functions are 1s with scaling factor $\alpha_A=\alpha_B=1.0$.

d) The same as a) but using A=W(C,[0,0,1],0,1.57).

e) From reference [71]: Equilateral triangular configuration of H_3 with R=2.0 au and using 1s functions with the scaling factor α=1.0.

Table 7.2

Integral type	Geometry[a]	Number of functions in L_n	Integral value
(AB,AB)	A - 3 - B	5	0.06037
		10	0.05889
		15	0.05855
		20	0.05560
(AB,BC)	A-1.5-B-1.5-C	5	0.24655
		10	0.24799
		15	0.24736
		20	0.24783
(AB,CD)	A-1-B-1-C-1-D	5	0.31044
		10	0.31129
		15	0.31857
		20	0.30847
(AA,BB)[b]	A - 3 - B		0.31980

a) All the geometries are linear. The numbers shown between centers indicate the intercenter distance expressed in atomic units.

b) Exact calculation using formula (7.25).

7.7. Unified Treatment of Some One and Two Electron Integrals

Nuclear attraction integrals on one hand, and repulsion integrals on the other hand, are the first elements of a general integral formulation connected with atomic and molecular properties as well as with relativistic corrections. A good review of all the possible related operators involved in this family can be found in the GTO framework within the work of Matsuoka [68a,b] and in the excellent textbook of McWeeny [72], if detailed operator information and physical significance are sought.

It is easy to see that many operators, containing negative powers of the distance, play a leading role in the mentioned fields. In both one and two electron cases there is a general operator pattern which can be written as:

$$F_w(\mathbf{v}_1,\mathbf{v}_2,\mathbf{w},k) = M_w(\mathbf{v}_1,\mathbf{v}_2,\mathbf{w}) \; | \mathbf{v}_1-\mathbf{v}_2 |^{-k}, \qquad (7.33)$$

where the vector pair $\{\mathbf{v}_1,\mathbf{v}_2\}$ is associated in the one electron case to the electron-nuclear positions vector couple $\{\mathbf{r}_1,\mathbf{C}\}$ or to the electron pair positions $\{\mathbf{r}_1,\mathbf{r}_2\}$ in the bielectronic integral case.

The moment part of the operator: $M_w(\mathbf{v}_1,\mathbf{v}_2,\mathbf{w})$ is defined as in equation (6.17):

$$M_w(\mathbf{v}_1,\mathbf{v}_2,\mathbf{w}) = \prod_i (v_{1i} - v_{2i})^{w_i}, \qquad (7.34)$$

with $w=\Sigma_i w_i$ and $\{v_{1i},v_{2i}\}$ being the i-th components of the involved position vectors.

In this manner the one and two electron integrals can be studied from a general point of view.

7.7.1. One-electron Integrals

Nuclear attraction, field and field gradient are related to operators of the kind (7.33) [67], which can be particularly written as:

$$F_c(\mathbf{C},\mathbf{c},k) = M_c(\mathbf{C},\mathbf{c})\, r_C^{-k}, \tag{7.35}$$

where $M_c(\mathbf{C},\mathbf{c})$ is the electric moment operator constructed exactly as in equation (6.17). The operators (7.35) can also be associated to a CETO function like $\chi(\mathbf{C},\mathbf{c},-k,0)$. From this optic, it is straightforward to see that the same cases as in nuclear attraction integrals, discussed early in section 6.6 do appear. In the nuclear attraction framework, the operator in equation (7.35) formalism may be written as $F_0(\mathbf{C},\mathbf{0},1)$. The field operators, taken as another example, will be located into the operator set $F_1(\mathbf{C},\mathbf{c},3)$.

The same formulae and procedures as those developed in the four divisions of section 6.6 can be applied substituting F_0 by the appropriate F_c operator.

A slightly different situation may be related to the electron spin-same orbit interaction which can be considered as a blend of F_1 type operators and gradient derivative elements.

7.7.2. Two-electron Integrals

The remarks of the preceding section can be also applied in the two electron case. Here the operator forms can be written:

$$G_w(1,2,w,k) = M_w(r_1,r_2,w) \; | r_1-r_2 |^{-k}, \tag{7.36}$$

When dealing with the reduced integrals the operator G_w can be integrated over the electron (2) coordinates in the same manner as has been discussed in section 7.2.1., obtaining a function formally defined as:

$$G(1) = \langle W(B,b,m,\beta) | G_w(1,2,w,k) \rangle \{1\}, \tag{7.37}$$

which is the generalization of equation (7.5) for this kind of operators. Like in equation (7.7) it can be written:

$$G(1) = \Sigma(i=0,b) \; \Sigma(j=0,i) \; M(i,j,b) \; f(i,j,b,\varphi_1,\theta_1) \cdot \langle W(B,g,m,\beta) | W(\{1\},w,-k,0) \rangle. \tag{7.38}$$

Finally, the general two electron reduced integral:

$$\langle A \mid G_w \mid B \rangle = \langle W(\mathbf{A},\mathbf{a},n,\alpha) \mid G(1) \rangle \qquad (7.39)$$

can be evaluated under similar considerations as these discussed previously when repulsion integrals where studied.

The same problems encountered when discussing Coulomb integrals may be found in this case, even worsened by the presence of higher negative powers of the distance. O-ohata and Ruedenberg transformation [26b] can be also used here, and thus the problem reduces to a first integration over an overlap-like expression, followed by a spherical integration of a potential-like function.

In this manner, practically all the most usual integrals needed in MO calculations are defined in terms of WO-CETO functions.

Conclusions and Prospects of Future

A general, simple and pedagogical framework based on CETO functions has been discussed in depth, in order to obtain atomic and molecular integrals, which can be potentially used within a LCAO computational system. One and two electron integrals over various operator kinds have been solved and analytical forms found as well, proving in this manner the flexibility and complete possibilities of the proposed methodology.

Two center density expansions into separate centers have been employed successfully in order to overcome the many center integral problem. Here it is also proved how such a naïve but elegant algorithm, based essentially on a recursive Cholesky decomposition, can be well adapted to modern computational hardware architectures. CETO functions appear in this manner as a plausible alternative to the present GTO quantum chemical computational flood, constituting the foundation of another step signaling the path towards STO integral calculation.

Many interesting methodological ideas, along the description of CETO's and their alternative forms: SETO's, LETO's, WO-CETO's and E-CETO's, have been also defined: Logical Kronecker delta and nested summation symbols among others. Both concepts permit easily written integrals and related formulae within a compact mathematical formalism which possess an immediate and intuitive translation to high level programming languages [45b]. From here the possibility to rewrite within this point of view other well known AO integral formulae appears promising.

Implementation of the whole set of integral algorithms within the ARIADNE molecular program [66a] as well as MOLSIMIL molecular Quantum Similarity code [66b], developed in our Laboratory is under way. A discussion on the sequential, vector and parallel programming features of the CETO integral calculation will be published elsewhere. Perhaps other available ETO functions, left unexplored on this paper, will be studied in the near future and the

application of the present formalism to GTO framework may also be tried. A huge amount of work is waiting to be done.

Acknowledgements

One of the authors (R.C.) wants to thank Prof. S. Fraga for the challenging proposal made to him in 1969, concerning computation of three center STO nuclear attraction integrals. The interest of Prof. S. Huzinaga on the possible use of STO's in the same manner as GTO's, and the test atomic calculations made by him in 1977 on the second and third rows of the periodic table [73] are also deeply acknowledged.

This work has been supported by the Spanish "Ministerio de Industria y Energía" under the "Programa Nacional de Investigación y Desarrollo Farmacéuticos" throughout a grant #: FAR 88-0617. One of us (E.B.) benefits of a grant of the Departament d'Ensenyament of the Generalitat de Catalunya.

References

[1] S.F.Boys and G.B.Cook; Revs.Mod.Phys. **32** 285 (1960)

[2] a) S.F.Boys; Proc.Roy.Soc. (London) **A200** 542 (1950)
b) S.Huzinaga, J.Andzelm, M.Klobukowski, E.Radzio-Andzelm, Y.Sakai and H.Tatewaki: "Gaussian Basis Sets for Molecular Calculations". Elsevier. Amsterdam (1984)
c) S.Huzinaga: "Approximate atomic functions". Technical reports I, II (1971) and III (1973) Division on Theoretical Chemistry. Department of Chemistry. University of Alberta.

[3] E.Hylleraas; Z.Physik **51** 150 (1928)

[4] E.C.Kemble and C.Zener; Phys.Rev. **33** 512 (1929)

[5] J.H.Bartlett; Phys.Rev. **37** 507 (1931)

[6] N.Rosen; Phys.Rev. **38** 255 (1931)

[7] a) J.O.Hirschfelder, H.Diamond and H.Eyring; J.Chem.Phys. **5** 695 (1935)
b) J.O.Hirschfelder; J.Chem.Phys. **6** 795 (1938)
c) J.O.Hirschfelder and C.N.Weygandt; J.Chem.Phys. **6** 806 (1938)

[8] a) C.A.Coulson; Proc.Cambr.Phil.Soc. **33** 104 (1937)
b) C.A.Coulson; Trans.Far.Soc. **33** 388 (1937)
c) C.A.Coulson; Proc.Cambr.Phil.Soc. **34** 204 (1938)
d) C.A.Coulson; Proc.Cambr.Phil.Soc. **38** 210 (1941)

[9] a) P.O.Löwdin; Arkiv Mat.Astr.Fys. **35A** numbers 9,1 (1947) and 30,1 (1948)
b) P.O.Löwdin: "A theoretical investigation into some properties of ionic crystals". Thesis; Almquist and Wikseis, Uppsala (1948)

[10] R.S.Mulliken, C.A.Rieke, D.Orloff and H.Orloff; J.Chem.Phys. **17** 1248 (1949)

[11] H.J.Kopineck; Z.Naturfursch **5a** 420 (1950)

[12] J.O.Hirschfelder and J.W.Linnett; J.Chem.Phys. **18** 130 (1950)

[13] S.O.Lundquist and P.O.Löwdin; Ark.Phys. **3** 147 (1951)

[14] a) K.Ruedenberg; J.Chem.Phys. **19** 1459 (1951)
b) K.Ruedenberg, C.C.J.Roothaan and W.Jaunzemis; J.Chem.Phys. **24** 201 (1956)

[15] M.P.Barnett and C.A.Coulson; Phil.Trans.Roy.Soc. (London) **243** 221 (1951)

[16] P.O.Löwdin; Adv.Phys. **5** 1 (1956)

[17] a) C.C.J.Roothaan; J.Chem.Phys. **19** 1445 (1951)
b) C.C.J.Roothaan; J.Chem.Phys. **24** 947 (1956)

[18] W.C.Hamilton; J.Chem.Phys. **26** 1018 (1957)

[19] I.Shavitt and M.Karplus; J.Chem.Phys. **36** 550 (1962)

[20] a) A.Lofthus; Mol.Phys. **5** 105 (1962)
b) A.Lofthus; Mol.Phys. **6** 115 (1963)

[21] F.J.Corbató and A.C.Switendick: "Methods in Computational Physics". Vol **2**. Academic Press Inc. New York (1963)

[22] a) F.E.Harris; Revs.Mol.Phys. **35** 558 (1963)
b) F.E.Harris and H.H.Michels; Adv.Chem.Phys. **13** 205 (1967)
c) F.E.Harris; J.Chem.Phys. **51** 4770 (1969)

[23] M.Kotani, A.Amemiya, E.Ichigura and T.Kimura: "Tables of Molecular Integrals". Maruyen. Tokyo (1963)

[24] R.M.Pitzer and W.N.Lipscomb; J.Chem.Phys. **39** 1995 (1963)

[25] S.Fraga; Can.J.Chem. **42** 2509 (1964)

[26] a) K.Ruedenberg, K.O-ohata and D.G.Wilson; J.Math.Phys. **7** 539 (1966)
b) K.O-ohata and K.Ruedenberg; J.Math.Phys. **7** 547 (1966)
c) R.E.Christoffersen and K.Ruedenberg; J.Chem.Phys. **49** 4285 (1968)
d) L.S.Salmon, F.W.Birss and K.Ruedenberg; J.Chem.Phys. **49** 4293 (1968)
e) D.M.Silver and K.Ruedenberg; J.Chem.Phys. **49** 4301 (1968)

[27] H.J.Silverstone; J.Chem.Phys. **48** 4106 (1968)

[28] P.O.Löwdin; Adv. Quantum Chem. **5** 185 (1970)

[29] S.Bosanac and M.Randić; J.Chem.Phys. **56** 337 (1972)

[30] J.L.Whitten; J.Chem.Phys. **58** 4496 (1973)

[31] R.R.Sharma; Phys.Rev. **A13** 517 (1976)

[32] C.A.Weatherford and B.L.Jain; Intl.J.Quantum Chem.Symp. **14** 443 (1980)

[33] C.A.Weatherford and H.W.Jones (Editors): "ETO Multicenter Molecular Integrals". D.Reidel Pub.Co. Dordrecht (1982)

[34] a) R.López, G.Ramírez, J.M.García de la Vega and J.Fernández Rico; J.Chim.Phys. **84** 695 (1987)
b) J.Fernández Rico, R.López and G.Ramírez; J.Comp.Chem. **9** 790 (1988)
c) J.Fernández Rico, R.López and G.Ramírez; J.Comp.Chem. **10** 869 (1989)
d) J.Fernández Rico, R.López and G.Ramírez; Int.J.Quantum Chem. **37** 69 (1989)
e) J.Fernández Rico, R.López and G.Ramírez; J.Chem.Phys. **91** 4204 and 4213 (1989)
f) J.Fernández Rico, R.López and G.Ramírez; J.Chem.Phys. **94** 5032 (1991)

[35] a) I.I.Guseinov; J.Phys.B: Atom.Molec.Phys. **3** 1399 (1970)
b) I.I.Guseinov; J.Chem.Phys. **65** 4718 and 4722 (1976)
c) I.I.Guseinov; J.Chem.Phys. **67** 3837 (1977)
d) I.I.Guseinov and F.S.Sadichov; J.Phys. B: Atom.Molec.Phys. **10** L261 (1977)

e) I.I.Guseinov and E.M.Imamov;
J.Phys. B: Atom.Molec.Phys. **11** L475 (1978)
f) I.I.Guseinov; J.Chem.Phys. **69** 4990 (1978)
g) I.I.Guseinov; J.Chem.Phys. **72** 2198 (1980)
h) I.I.Guseinov; Phys.Rev. A **22** 369 (1980)
i) I.I.Guseinov; in reference [33] pg. 171.
j) I.I.Guseinov; J.Chem.Phys. **78** 2801 (1983)
k) I.I.Guseinov; Phys.Rev. **A31** 2851 (1985)
l) I.I.Guseinov;
Intl.J.Quantum Chem.(Q.C.Symp.) **19** 149 (1986)
m) I.I.Guseinov in: "Quantum Chemistry: Basic Aspects, Actual Trends" R.Carbó (Editor).
Studies in Physical and Theoretical Chemistry.
Vol **62**. Pg. 37. Elsevier. Amsterdam (1989)

[36] a) H.W.Jones; Int.J.Quantum Chem. **20** 1217 (1981)
b) H.W.Jones; Int.J.Quantum Chem. **21** 1079 (1982)
c) H.W.Jones; Phys.Rev. **A30** 1 (1984)
d) H.W.Jones; Int.J.Quantum Chem. **29** 177 (1986)
e) H.W.Jones, B. Bussery and C.A.Weatherford;
Int.J.Quantum Chem.(Q.C.Symp.) **21** 693 (1987)
f) H.W.Jones; Phys.Rev. **A38** 1065 (1988)
g) H.W.Jones and B.Etemadi;
Int.J.Quantum Chem.(Q.C.Symp.) **24** 405 (1990)

[37] a) See for example, E.O.Steinborn in ref [33], Pg. 7 and references therein.
b) E.O.Steinborn in: "Methods in Computational Molecular Physics". G.H.F. Diercksen et. al. (Editors). Pg. 37. D.Reidel Pub. Co. Dordrecht (1983)
c) E.O.Steinborn and H.H.H.Homeier;
Intl.J.Quantum Chem.(Q.C.Symp.) **24** 349 (1990)
d) E.O.Steinborn and E.J.Weniger;
J.Mol.Struc.(Theochem) **210** 71 (1990)
e) H.H.H.Homeier and E.O.Steinborn;
Intl.J.Quantum Chem. **39** 625 (1991)

[38] K.Ruedenberg; J.Chem.Phys. **19** 1433 (1951)

[39] R.Carbó and J.M.Riera: "A General SCF Theory" Lecture Notes in Chemistry. Vol. 5 Springer Verlag. Berlin (1978)

[40] a) R.Carbó and C.Arnau; Gazz.Chim.Ital. **108** 171 (1978)
b) R.Carbó and C.Arnau; Intl.J.Quantum Chem. **14** 209 (1978)
c) R.Carbó and C.Arnau; Afinidad **35** 171 (1978)

[41] R.Carbó and E.Besalú; Can.J.Chem. ?? ??? (199?)

[42] V.R.Saunders in: "Computational Techniques in Quantum Chemistry and Molecular Physics". G.H.F.Diercksen et al. (Editors). Pg. 347. D. Reidel Pub. Co. Dordrecht (1975)

[43] a) J.C.Slater; Phys.Rev. **36** 51 (1930)
b) E.Clementi and C.Roetti; At.Nucl.Data Tab. **14** 177 (1974)
c) see reference [2c]

[44] R.Carbó and B.Calabuig in: "Concepts and Applications of Molecular Similarity". Chapter 6. Pg. 147. M.A.Johnson and G.M.Maggiora (Editors). John Wiley & Sons Inc. New York (1990)

[45] a) R.Carbó and J.A.Hernández; Chem.Phys.Lett. **47** 85 (1977)
b) E.Besalú and R.Carbó; Computers and Chemistry (submitted)
c) R.Carbó and C.Bunge; PC Actual Magazine. September 1989. Pg. 124.

[46] P.W.Atkins: "Quanta". Oxford Chemistry Series. Clarendon Press. Oxford (1974)

[47] A.C.Wahl, P.E.Cade and C.C.J.Roothaan; J.Chem.Phys. **41** 2578 (1964)

[48] a) A.D.McLean and Y.S.Lee in: "Current Aspects of Quantum Chemistry 1981". Pg. 219. R.Carbó (Editor) Studies in Physical and Theoretical Chemistry **21**. Elsevier Pub. Co. Amsterdam (1982)
b) E.Clementi (Editor): "Modern Techniques in Computational Chemistry 1990". IBM Corporation. Kingston (1990)

[49] a) C.G.Gray and K.E.Gubbins: "Theory of Molecular fluids". Vol. 1, appendix A. Clarendon Press. Oxford (1984)
b) R.N.Zare: "Angular Momentum". John Wiley & Sons Inc. New York (1988)

[50] a) W.Gröbner and N.Hofreiter: "Integraltafel". Springer-Verlag. Wien (1966)
b) M.R.Spiegel: "Mathematical Handbook of Formulas and tables". McGraw-Hill, Inc. New York (1968)
c) M.Abramowitz and I.A.Stegun: "Tables of Mathematical functions". Dover Pub. Inc. New York (1972)

[51] a) 3L Ltd.: "Parallel Fortran" Livingston. Scotland (1988)
b) INMOS Ltd.: "Transputer Reference Manual". Prentice Hall. New York (1988)
c) INMOS Ltd.: "Transputer Technical Notes" Prentice Hall. New York (1989)

[52] a) G.Rodrigue (Editor): "Parallel Processing for Scientific Computing". SIAM. Philadelphia (1989)
b) K.A.Galivan et al.: "Parallel Algorithms for Matrix Computations". SIAM. Philadelphia (1990)
c) A.Wouk (Editor): "New Computing Environments: Parallel, Vector and Systolic". SIAM. Philadelphia (1986)
d) A.Wouk (Editor): "Parallel Processing and Medium-Scale Multiprocessors". SIAM. Philadelphia (1989)

[53] J.A.Pople and D.L.Beveridge: "Approximate MO Theory". McGraw Hill. New York (1970)

[54] a) C. Zener and V.Guillemin; Phys.Rev. **34** 999 (1929)
b) W.M.De Jeu; Chem.Phys.Lett. **5** 109 (1970)

[55] a) E.Harris and R.Rein; Theoret.Chim.Acta **6** 73 (1966)
b) M.D.Newton; J.Chem.Phys **51** 3917 (1969)
c) A.Okninski; J.Chem.Phys **60** 4098 (1974)

[56] a) J.H.Wilkinson: "The Algebraic Eigenvalue Problem". Clarendon Press. Oxford (1965)
b) E.Durand: "Solutions Numériques des Equations Algebriques" Masson et Cie. Paris (1961)
c) R.Carbó and Ll.Domingo: "Algebra Matricial y Lineal" Serie Schaum. McGraw Hill. Madrid (1987)
d) J.H.Wilkinson and C.Reinsch: "Linear Algebra" Springer Verlag. Berlin (1971)

[57] K.O-ohata,H.Taketa and S.Huzinaga; J.Phys.Soc. Japan **21** 2306 (1966)

[58] G.S.G.Beveridge and R.S.Schechter: "Optimization: Theory and Practice" McGraw-Hill Kogakusha, Ltd. Tokyo (1970)

[59] a) Lahey Computer Systems Inc.: "Fortran Language F77L/EM32". Version 3.10. Revision A. June 1990. Incline Village. Nevada.
b) Lahey Computer Systems Inc.: "Fortran Language F77L". Version 4.10. Revision E. August 1989. Incline Village. Nevada.

[60] MicroWay Inc.: "NDP Fortran-386". Version 3.0. December 1990. Kingston. Massachussets.

[61] G.Coshi and J.B.Schueler: "WATCOM Fortran 77. Language Reference". WATCOM Pub. Ltd. Waterloo. Ontario (1990)

[62] K.W.Christopher Jr.,B.A.Feigenbaum and Sh.O.Saliga: "The New DOS 4.0". John Wiley & Sons Inc. New York (1989)

[63] R.Carbó and B.Calabuig in: "Quantum Chemistry. Basic Aspects, Actual Trends". R.Carbó (Editor). Studies on Physical and Theoretical Chemistry. Vol. **62**, pg. 73. Elsevier. Amsterdam (1989)

[64] Programmer's Connection Inc.: The Connection **5** #2. North Canton. Ohio (1991)

[65] a) H.Tai; NASA Technical Memorandum 101545, Hampton, Virginia 23665-5225 (1989)
b) H.Tai; Phys. Rev. A. **40** 6681 (1989)
c) H.Tai; Chin.J.Phys. **28** 481 (1990)

[66] a) R.Carbó and B.Calabuig; Comput.Phys.Commun. **52** 345 (1989)
b) R.Carbó and B.Calabuig; Comput.Phys.Commun. **55** 117 (1989)

c) R.Carbó and Ll.Domingo;
Intl.J.Quantum Chem. **32** 517 (1987)
d) R.Carbó and B.Calabuig;
Intl.J.Quantum Chem. ?? ??? (199?)

[67] V.R.Saunders in: "Methods in Computational Molecular Physics". G.H.F.Diercksen and S.Wilson (Editors). Pg. 1. D.Reidel Pub. Co. Dordrecht (1983)

[68] a) O.Matsuoka; Intl.J.Quantum Chem. **5** 1 (1971)
b) O.Matsuoka; Intl.J.Quantum Chem. **7** 365 (1973)

[69] H.Taketa, S.Huzinaga and K.O-ohata;
J.Phys.Soc. Japan **21** 2313 (1966)

[70] A.C.Hurley; Reprint of Rev.Mod.Phys. **32** 400 (1960) in Quantum Theory of Molecular Electronic Structure. R.G.Parr. Benjamin, New York (1963)

[71] J.O.Hirschfelder and C.N.Weygandt; J.Chem.Phys. 806 **6** (1938)

[72] R.McWeeny: "Methods of Molecular Quantum Mechanics". Academic Press. London (1989)

[73] S.Huzinaga; J.Chem.Phys. **67** 5973 (1977)

EQUIVALENCE OF MATHEMATICAL OBJECTS OF INTEREST IN CHEMISTRY AND PHYSICS

Sherif El-Basil[(a)] and Milan Randić[(b)]

(a) Faculty of Pharmacy, Kasr El-Aini Str., Cairo 11562, Egypt.

(b) Department of Mathematics and Computer Science, Drake University, Des Moines, Iowa 50311, U.S.A.

Table of Contents

ABSTRACT

We review various aspects of equivalence relations occuring in chemistry and physics which are of potential interest to chemists. In particular we illustrate various equivalences that yield alternative representation of a same problem or that lead to novel insights. The

review focuses on the use of graphical graph invariants as well as graph codes, graph algorithms, and graph characterizations to illustrate: how viewing "old things from a new point of view" may lead to novel results, how old results can be obtained in a simpler way, how novel concepts emerge, and how hidden results become apparent. The review is divided into three parts, the first gives a general overview of the subject, while the second part considers a particular application to benzenoid compounds illustrating data reduction aspects of use for equivalences. The third part deals with the fractal dimensions of graphs and correlation with physical properties.

> "The formulation is mathematically equivalent to more usual formulations. There are therefore, no fundamentally new results. However, there is a pleasure in recognizing old things from a new point of view".
>
> Feynman

PART I

INTRODUCTION

1.1 MODELS AND MODELING

Mathematical modeling is concerned with representing the essential features of a selected physico-chemical reality by mathematical objects, that is by "translating" physics and chemistry of a system of interest into expressions which allow some mathematical manipulations. Subsequently one then hopes to answer proposed questions confirm or reject a working hypothesis, illustrate theorems or support conjectures which may lead to new mechanisms or novel experiments. Modeling is predominantly a theoretical endeavour, but equally modeling can be experimental since it allow one to use physical objects such as various building blocks, e.g. "Sticks and balls" representation of atoms and bonds. Today some such "experimental" modeling is accomplished with the help of computers, using various simulations, computer graphics, etc. The double helix model for DNA was concluded using standard molecular models.

Some models may appear very different from the objects they represent. Such "abstract" models may be rewarding because they are likely to reveal qualities of a system that may not even be suspected. There are numerous worthy models and modeling in chemistry relating to structures, reactions, transition states, spectra, bulk or local properties of large molecules, etc. Fig. 1 shows how a standard organic chemistry text [1] "models" the molecule of benz[a]anthracene and various other modelings of this benzenoid system at various levels of abstaction.

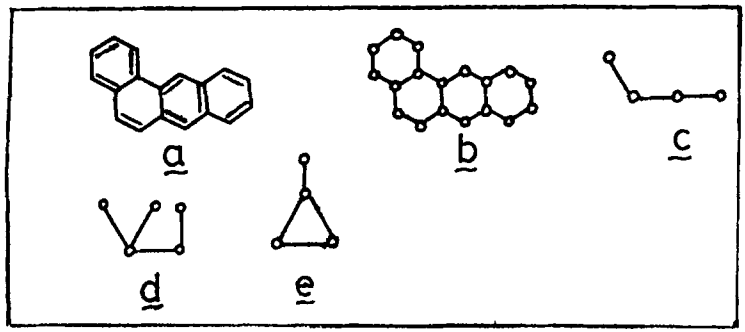


Fig. 1 Various modelings of benz[a]anthracene.

The mathematical chemist will recognize a as one of the Kekulé valence-bond structures of the hydrocarbon, while b is the corresponding molecular graph. Model c is called an inner dual or dualist [2], d is a caterpillar tree [3], and e is called a Clar graph [4]. The latter two models are apparently quite different from the original skeleton, however, as it will turn out later, the topological properties of this benzenoid system are best modeled by either d or e.

Instead of "graphical" modeling one may use another mathematical formulation, e.g., the benzenoid hydrocarbon shown in Fig. 1 may be described using a sequence of L,A symbols introduced by Gutman [5] in this case LAL^2. Alternatively Balaban [6] adopts a binary code.

Therefore graphs are not the only objects which can model systems, other tools are also available. Examples are shown in Table 1, where the first column represents the traditional more familiar ones while the second column lists a number of less familiar objects, still of interest in chemistry.

Table 1 Mathematical objects

Numbers	Tableau
Ordered Pairs	Lattices
Sequences	Partial orders
Functions	Configurations
Polynomials	Knots
Determinants	Partitions
Matrices	
Distributions	
Sets	
Groups	
Graphs	
Polyhedra	

One of the present authors [7] used caterpillar trees to model Wreath-product groups while Balasubramanian [8] used a particle-in-a-box model to represent the same type of permutation.

In considering graphs as objects in chemistry and related disciplines we will focus attention to problems in which alternative models or representations are possible in order to illustrate equivalence and its uses in chemical modeling.

1.2 EQUIVALENCE

A relation R is an equivalence relation [9] if it is reflexive, symmetric and transitive. A binary relation may possess all three properties, it may possess one, two or none of these properties as is illustrated below:

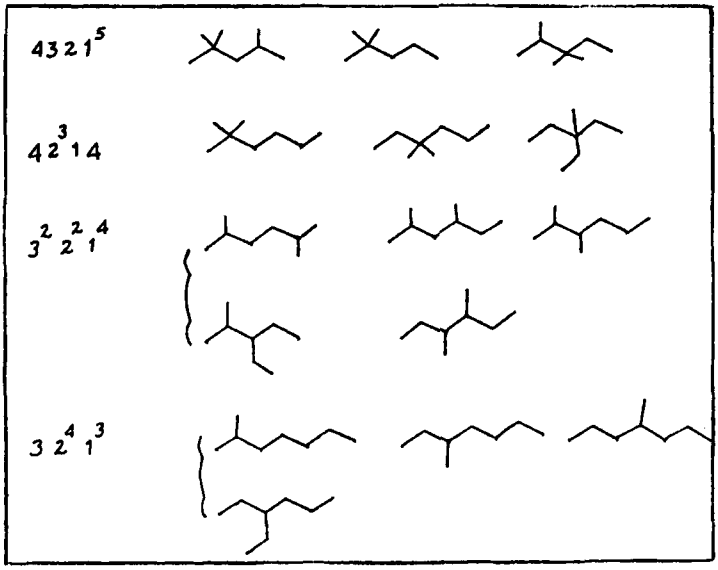
RELATION	REFLEXIVE	SYMMETRIC	TRANSITIVE
"is not equal to"		x	
"is greater than"			x
"is divisable by"	x		x
"is parallel to"	x	x	x
"is perpendicular to"		x	
"is similar to"	x	x	x
"is isomorphic"	x	x	x
"is equal to"	x	x	x
"is equivalent to"	x	x	x

Equivalence Classes: Any two elements of a set that are related to each other by an equivalence relation are said to be equivalent. All the elements of a set equivalent to a given element form an equivalence class.

In Table 2 we illustrate equivalence classes for some selected sets modeled by alkane molecular graphs [10].

"Equivalence" is also used as a synonym for "equality with respect to". This phrase implies that two such equivalent objects are not equal in all respects, but equal with respect to a property of interest. For example, the well known aromaticity $4n+2$ rule classifies even-polycyclic conjugated benzenoids into two classes. Under such a relation molecules such as naphthalene and anthracene are considered equivalent.

Table 2 Equivalence classes. Sequences to the left are vertex-degrees listed in a descending order [10].



1.3 GRAPHS

A graph is defined as a set of elements V, with a binary relation E defined on the set. Usually circles represent vertices (elements) and lines, usually referred to as edges, represent binary relations. The binary relation is both symmetric and antireflexive. The classical text of Harary's Graph Theory [11] collects several of the more popular types of graphs. One observes the inherent generality of the definition of graphs: There are no restrictions on what the set of elements, V, represents or to what the binary relation, E, corresponds. If we model a chemical structure, V, may be the set of atoms, while E the set of bonds present, but equally E may represent "close" atomic contacts (such as those occuring in crowded molecules, etc.). Atlernatively V can represent the set of Kekulé valence structures in a polycyclic conjugated hydrocarbon and E the relationship representing permutation of CC double and single bonds within a single benzene ring. If we consider the four Kekulé structures of anthracene we generate a complete graph, K_4, Fig. 2.

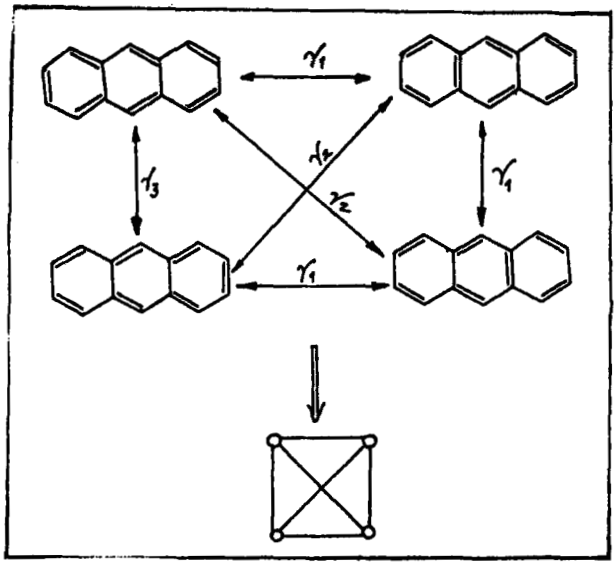


Fig. 2: Graphical modeling of the generation of permutation integrals of Herndon [12] structure-resonance theory in case of anthracene γ_i involves permutation of (4i + 2), pi-electrons.

Graphical modeling can also be useful in representing the elements of the transfer matrix, T, adopted by Klein et al., [13] Fig. 3 shows the five Kekulé valence-bond structures of phenanthrene and their local states. In this case one needs a directed graph with weighted edges and loops.

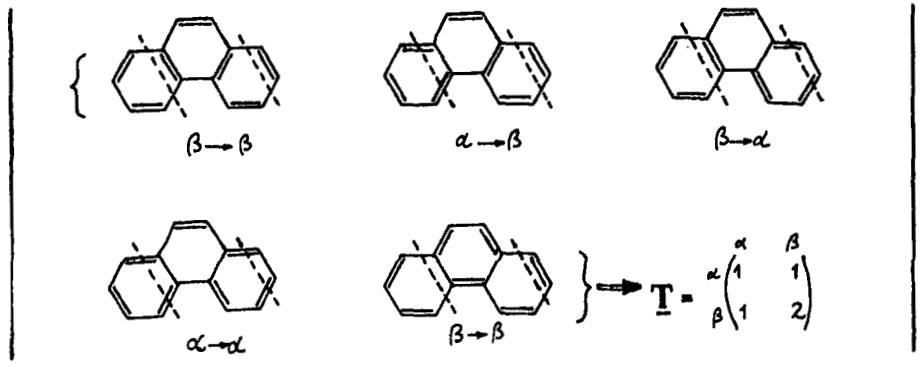


Fig. 3: Generation of the elements of the transfer matrix, T, corresponding to the five local states of phenanthrene.

Still another application V can represent a set of CC double bonds within a single Kekulé structure while E represents "adjacent" CC double bonds; The resulting graph is known as factor graph [14]. The relation between the later and a set of Kekulé structures is not bijective [15] as illustrated in Fig. 4.

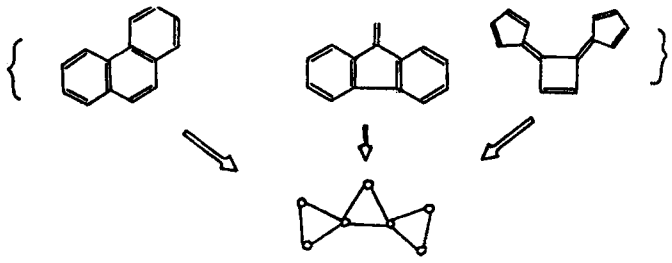


Fig. 4: Three Kekulé structures all modeled by the same factor graph.

When the hexagons of a benzenoid hydrocarbon are replaced by vertices and then connection is made between any two vertices corresponding to two nonresonant hexagons, then the resulting graph is called a Clar graph [4] (c.f. e in Fig. 1). Furthermore, if the benzenoid system is unbranched [16] its Clar graph turns out to be the line graph [17] of a caterpillar tree. Whence both caterpillar and Clar graphs are sophisticated graphical representations of benzenoid systems. In Fig. 1 e is the line graph of d. The derived simplified graphs need not only serve as compact notation but may allow novel manipulations leading to simpler codes, easier enumerations, exhaustive constructions etc.

There are two kinds of equivalences of interest when considering graphs:

(i) Search for alternative forms of a given graph;
(ii) Search for alternative objects, which parallel selected properties of graphs.

The first aspect relates to the graph isomorphism problem, but instead of an emphasis on verifying that two graphs G_1 and G_2 have identical connectivity we start with G_1 and search for alternative form of it which displays some of its inherent properties not apparent from a given representation. The second aspect will be illustrated in Part II, in which we show that a tree, a cyclic graph, a polyhex and a fragment of a board are all equivalent with respect to many selected properties. Such set of graphs can be called "equinumerical", because they lead to the same count of the selected property. In Fig. 5 we illustrate a number of equinumerocities the cases of a count of different objects which gives the same outcome. Thus, for example, the count of all canonical spin singlet valence structures (i.e., the count of possible Rumer diagrams) can be brought into one-to-one

correspondence with the construction of different spin functions, and this parallels the different ways of arranging brackets (corresponding to alternative ways of nested loops in programming). This is equivalent to the number of ways to parenthesize a product by means of inserting enough parentheses so that every subproduct is the multiplication of exactly two factors. As can be seen, this is also equivalent to counting different ways of arranging nonintersecting circles which is equivalent to enumerating the number of ways of triangulation of a pentagon [18]. The same count also enumerates the number of rooted trees on five vertices. In all these cases the resulting count is 5 which is the third member of the Catalan numbers [18], 1,2,5,14,42,132, ... The importance of equinumerocity is not only in the recognition of

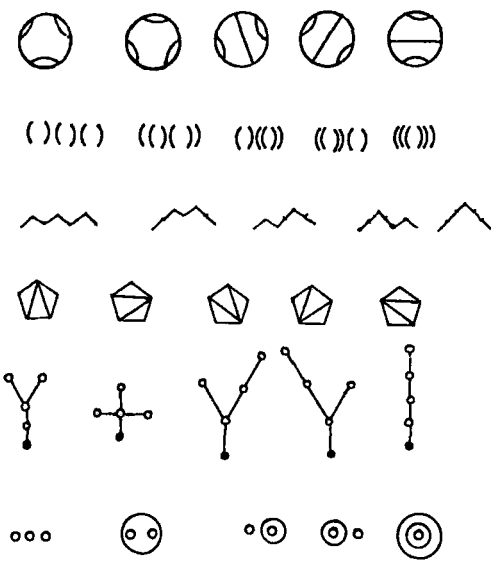


Fig. 5: Various ways of generating third catalan number.

mathematical equivalence of apparently different problems, but also in the practical aspect in that some of the enumeration may be easier to perform than others.

In this review we will focus on polyhex graphs, caterpillar trees, Clar graphs and several related polyomino graphs [19]. In addition sets of graphs obeying certain types of recursive relations, called "Fibonacci Graphs" will be discussed particularly from the point of view of their computational importance.

1.4 CHEMICAL GRAPH THEORY

Characterization of graphs and the study of their properties which are of interest in chemistry constitute chemical graph theory. The two volumes of Trinajstić [20] introduce the reader to the main domains of this branch of mathematical chemistry. The main graphs of chemical interest include Hückel graphs, benzenoid (polyhex) graphs, Möbius gaphs, Clar graphs, Fibonacci graphs and several forms of polyominos. Besides graphs, chemical graph theory is concerned with studying certain objects such as Kekulé structures, conjugated circuits [21], pi-electron permutation integrals and maximal independent sets of vertices [22] generated from certain types of colored graphs. Usually a gaph invariant [23] (such as nonadjacent numbers, [24] path numbers [25], random walks [26] etc.) are enumerated and a polynominal is constructed. A general form of a counting plynomial is given for a graph G on n vertices by [27]

$$F(G; X) = \sum_{k=0}^{M} \rho \, \theta(G; k) \, X^{f(k,n)}$$

where ρ is either 1 or $(-)^k$; $\theta(G; k)$ enumerates certain selected invariants in G taken in k independent tuples (i.e. no two are adjacent) and M is the maximal value k. A polynomial is defined in association to a particular graph so that the combinatorial function $\theta(G; k)$ lists certain k selections of particular graph invariants. Table 3 outlines some of the more important polynomials which have been in use in chemical graph theory.

The associated graphs of Table 3 act like "substrates" for which the polynomials were originally designed. In general the function $\theta(G; 0)$ is conveniently taken to be unity. When dealing with acyclic graphs, then the acyclic (matching) and characteristic polynomials coincide. The latter being derivable from diagonalization of the corresponding adjacency matrix, $\underline{A}$, and is defined as det $(\underline{A} - x\underline{I})$, where $\underline{I}$ is the unit matrix of the same dimension as $\underline{A}$.

Arbitrariness. One hears comments that graph descriptors being arbitrary, do not have a deep physicochemical significance. A misconception here is in confusing an input with an output of a mathematical treatment of a problem. Surely the choice of descriptors selected to characterize graphs is at the disposal of an investigator, hence arbitary. But the same is true of a choice of coordinate systems used to describe a system of classical physics, or the choice of basis functions to compute a molecular structure in quantum chemistry! In each case we speak of input information, and the outcome of the analysis undertaken (when a complete basis is taken) does not depend on the choice of descriptors (coordinates). However, the amount of work, even its mere feasibility, will differ. A poor selection of basis functions may

Table 3: Some Polynomials of Mathematical Chemistry.

	Polynomial	ρ	Graph invariants set	Associated graph	$\Theta(G;k)$	$f(k,n)$
1	Acyclic = matching $\equiv \alpha(G;x)$	$(-1)^k$	edges	caterpillar tree, T	$p(G;k)$[a]	$n-2k$
1(a)	Counting (of Hosoya) $\equiv H(G;x)$	1	edges	caterpillar tree, T	$p(G;k)$[a]	k
2	Sextet $\equiv \sigma(B;x)$	1	hexagons	polyhex graph, B	$r(B;k)$[b]	k
2(a)	Resonance $A(B;x)$	$(-1)^k$	hexagons	polyhex graph, B	$r(B;k)$[b]	$2M-2k$
3	King $\equiv K(P;x)$	1	cells	polyomino graph, P	$K(P;k)$[c]	k
3(a)	Rook $\equiv K(P_r;x)$	1 or $(-1)^k$	cells	polyomino graph, P	$\rho(P_r;k)$[d]	k or $n-2k$
4	Independence $\omega(\Lambda;x)$	1	vertices	Clar graph, Λ	$0(\Lambda;k)$[e]	k
4(a)	Color $C(G;x)$	1	vertices	arbitrary graph, G	$\xi(G;k)$[f]	k

[a] Number of selections of k independent edges $\in T$ (i.e. no two edges are incident).
[b] Number of selections of k nonadjacent but mutually resonant hexagons B.
[c] Number of ways of arranging k non-taking kings on a polyomino graph.
[d] Number of ways of arranging k non-attacking kings.
[e] Number of selections of k independent vertices $\in \Lambda$ (no two are adjacent).
[f] Number of colorings in G in which there are k vertices of the same color so that no two of them are adjacent.

lead to difficult (or even impossible to solve) molecular integrals, an inadequate selection of coordinates may make integration tedious. In a similar manner a poor selection of graph invariants will obscure regularities that would otherwise be easy to recognize. Regularities, as will be seen in Part II, are an important feature of sets of equivalent graphs. Such regularities make possible the computation of some characters of very large graphs only from knowledge of much smaller leading members.

1.5 GRAPH THEORETICAL ANALYSIS

In developing mathematical modeling in chemistry and physics various branches of mathematics have been widely employed, frequently each with a well defined domain of applicability. This sometimes has not been sufficiently emphasized or recognized, particularly when less familiar or less frequently used mathematical disciplines are involved, such as combinatorics, topology, graph theory, set theory or category theory. The lack of appreciation of the specifics of each such discipline

is sometimes reflected by people expecting answers for inappropriate questions. For example, referring to group theory one can answer questions concerning selection rules, symmetry, or pattern of splitting of degeneracy (under perturbation). However, questions relating to magnitudes of such splittings are often inappropriate, being outside the domain of group theory, as is today widely understood. The mission of graph theory, on the other hand, is to discern regularities of a given property in a set of structures. That is: given a structure and given a property, can we identify important structural factors responsible for the property considered? But to try to use graph theory to predict a property for a given structure without giving additional information on some related structures reflects unfamiliarity with the specific nature of chemical graph theory, as presently realized. Prediction of the properties of a single isolated molecule is, at least in principle, the domain of quantum mechanics, not graph theory. Should one have a sufficiently accurate wave function for the system, properties of interest could,in principle, be calculated. If one applies such methodology to, say, a dozen molecules one can derive the corresponding dozen results for the property considered. But that is where computational quantum chemistry usually ends. Most chemists would like to continue and try to see some "pattern" in the computed data. This amounts to reducing a lot of data to some simpler picture. Those efforts in data reduction, even though often not explicitly indicated or recognized as such represent some graph theoretical analysis. The work of Gutman on the representation of unbranched and branched polyhex graphs by caterpillars and Clar graphs, respectively, is an explicit form of such data reduction [28]. Another endeavour of graph theory is in ordering [29] problems. Randić's pioneering work in this area [29] presents, perhaps one of the elegant uses of graph theory in the analysis of the properties of isomeric alkanes using a two-dimensional structure space diagram where each isomer is located by p 2 and p 3 (numbers of paths of lengths 2 and 3 respectively). In Part II we will show how to use equivalence relations to rigorously order a wide variety of graph types based on partial ordering relations defined by Ruch and Schönhofer [30,31].

Alternative Forms

Sometimes using an alternative form of a graph may lead to the discovery of a regularity of a particular property that would not have been so transparent from the original model. A rather classical case is related to the description of localized orbitals: as a starting point we consider all honogenous functions $x^{\alpha} y^{\beta} z^{\gamma}$ of degree ℓ which satisfy Laplace's eqn. [32]. When $\ell = 0$ then $\alpha = \beta = \gamma = 0$ and the desired function is unity. For $\ell = 1$ there are three functions which are homogenous of degree 1 and satisfy Laplace's eqn., namely:

$$\alpha = 1,\ \beta = \gamma = 0 \longrightarrow u_1 = x$$
$$\alpha = 0,\ \beta = 1,\ \gamma = 0 \longrightarrow u_1 = y$$
$$\alpha = 0,\ \beta = 0,\ \gamma = 1 \longrightarrow u_1 = z$$

When the cartesian coordinates are transformed into spherical polar coordinates and normalization is taken into consideration, the three forms of p-orbitals take the following expressions.

$$P_x = \tfrac{1}{2}(\tfrac{3}{\pi})^{\frac{1}{2}} \sin\theta \cos\varphi \quad ; \quad P_y = \tfrac{1}{2}(\tfrac{3}{\pi})^{\frac{1}{2}} \sin\theta \sin\varphi$$
$$; \ P_z = \tfrac{1}{2}(\tfrac{3}{\pi})^{\frac{1}{2}} \cos\theta$$

when $\ell = 2$ one obtains six homogeneous functions namely:

$$\alpha = 2,\ \beta = \gamma = 0 \longrightarrow x^2$$
$$\alpha = 0,\ \beta = 2,\ \gamma = 0 \longrightarrow y^2$$
$$\alpha = 0,\ \beta = 0,\ \gamma = 2 \longrightarrow z^2$$
$$\alpha = \beta = 1,\ \gamma = 0 \longrightarrow xy$$
$$\alpha = \gamma = 1,\ \beta = 0 \longrightarrow xz$$
$$\alpha = 0,\ \beta = \gamma = 1 \longrightarrow yz$$

Further the six derived expressions are not linearly independent. Infact, there are only $2\ell + 1 = 5$ linearly independent functions and we must use the first three functions to construct two more linearly independent solutions to Laplace's eqn. We find that the functions $(x^2 - y^2)$, $(x^2 - z^2)$ and $(y^2 - z^2)$ do satisfy the desired condition. Arbitrarily choosing $(x^2 - z^2) + (y^2 - z^2) = r^2 - 3z^2$, where, of course, $x^2 + y^2 + z^2 = r^2$. Then one may obtain the five d orbitals, (c.f. Fig. 5), usually designated as d_{z^2}, d_{xz}, d_{yz}, $d_{x^2-y^2}$, d_{xy}. Similarly there are 10 homogenous functions of degree 3, but there are three conditions which reduce the degree of the functions, thus leading to $10 - 3 = 7$ f orbitals. While the above algebra is simple, similar analysis for equivalent expressions would have been tedious using polar coordinates. This illustrates the advantage of using an alternative model of a given system. On the other hand solving the Schrödinger eqn. for the hydrogen atom requires, for the separation of variables the transformation of cartesian coordinates into their spherical polar alternative forms.

A more striking case which illustrates the advantage of the equivalent alternative forms was discovered sometime ago by one of the authors [33]: An alternative form of a Kekulé valence-bond structure, K, is its factor graph, F(K), c.f. Fig. 4. In Fig. 6 all Kekulé structures of picene are shown (in the form of their Clar notation) along with the corresponding factor graphs. One recalls that an F(K) is

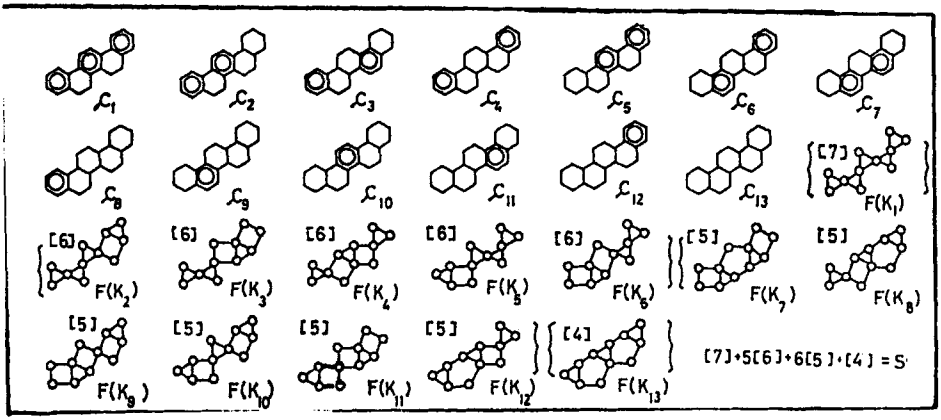


Fig. 6: Clar structures and factor graphs of picene.

constructed from a given K by repalcing each double bond by a vertex and then connecting any two vertices which correspond to two conjugated double bonds. The number of bivalent vertices in each F(K) is placed in square brackets and a partition sequence, S, is computed for each hydrocarbon. For picene El-Basil obtained [34].

$$S = [7] + 5[6] + 6[5] + [4]$$

which means that the set of F(K)'s of picene contains one F(K) possessing 7 bivalent vertices, 5 F (K)'s possessing 6 bivalent vertices each and so on. Now the sextet polynomial, σ, of picene is given by:

$$\sigma = 1 + 5x + 6x^2 + x^3$$

There is an identity between the coefficients of σ and the populations of the F(K)'s. El-Basil studied several classes of unbranched benzenoids and demonstrated several graphical identities. Such findings were clear only when modeling Kekulé structures as their equivalent factor graphs. Table 4 outlines some results. A true story that is worthy of mention is now noted to illustrate the advantage of using alternative forms: Several years ago Gutman [35] wrote a review on theorems, conjectures and unsolved problems related to the topological properties of benzenoid systems. El-Basil had troubles understanding what Gutman wrote as Theorem 42 in his review [35], namely.

$$a_k = \binom{h+1-k}{k}$$

where a_k is the number of Kekulé structures of an unbranched benzenoid hydrocarbon in which there are exactly 2K double bonds of the

Table 4: Partition sequence (above) and sextet polynomial (below) for the zigzag benzenoids.

R	Example	
3		[5]+3[4]+[3] $1+3X+X^2$
4		[6]+4[5]+3[4] $1+4X+3X^2$
5		[7]+5[6]+6[5]+[4] $1+5X+6X^2+X^3$
6		[8]+6[7]+10[6]+4[5] $1+6X+10X^2+4X^3$
7		[9]+7[8]+15[7]+10[6]+[5] $1+7X+15X^2+10X^3+X^4$
10		[12]+10[11]+36[10]+56[9] +35[8]+6[7] $1+10X+36X^2+56X^3+35X^4$ $+6X^5$

type C=CH. Gutman gave two references for his Theorem 42: both references were El-Basil's original papers !! To that extent, then, using another equivalent model could obscure one's very own results!

PART II

In this part we review three types of equivalencies viz, Graphs possessing equivalent alternative models, homologous series of graphs and a recently introduced [36] class of equivalent graphs of computational importance called Fibonacci graphs.

2-1. EQUIVALENT ALTERNATIVE MODELS

First we will focus attention on selected topics relating to the equivalence between benzenoid hydrocarbons, and special types of graphs and other mathematical objects that we can associate with benzenoids. In particular we will explore relations involving caterpillar trees [3] associated with catacondensed benzenoids and their line graphs [17] called, as already mentioned, Clar graphs [4]. Also relations involving "boards" (known technically as polyominos) of special properties such as those associated with "king" and "rook" pieces of chess

board and other related constructions will be mentioned and exemplified. The underlying thought in all such comparisons is that some such related objects may better display some of the properties of interest of the benzenoids, which are our prime interest. But as will be seen the "equivalence" extends over the traditional boundaries of chemistry, covering not only interdisciplinary fields of organic chemistry, with those of quantum chemistry and statistical physics, but even some areas of "recreational" mathematics. A common interest in many such apparently unrelated fields is in enumeration of certain qualified components. We have already seen in Fig. 5 that enumeration of apparently unrelated objects such as "nested loops", or "non-intersecting circles", or counting rooted trees, etc., are in one-to-one relationship, hence "solving" one such proplem gives the solution for all other equi-numerical tasks.

2.1.1 Benzenoids

Before we proceed, a word on benzenoids which are the central object of our interest here, seems in place. To focus enquiry on a family of some such specially selected structures of chemistry may appear as a very "narrow" goal. In addition some may add that even within the chemistry of hydrocarbons, benzenoids are a very select subgroup of structures so that any information so deduced may be of little interest outside the particular domain. However, as will be seen, such objections are not sound, they simply ignore equivalences among different chemical problems. In fact, by focusing attention to such equivalence relations one can obtain a more unified picture of chemistry (and parts of physics), than would follow from simple accumulation of the same such data but not recognizing the "common" traits in diverse problems. Clearly there is a considerable synergetic effect and as already demonstrated such beneficial influence of apparently unrelated problems not only makes easy or possible the resolution of some problems that appear untractable but add to our conceptualization of chemistry. In Table 5 we briefly summarize various aspects of the chemistry of benzenoids by listing a number of "key" words which appear in publications dealing with benzenoids.

Table 5: Some "key" terms in the literature of benzenoid hydrocarbons.

	Mathematical	Chemical
1)	Perfect matching	Kekule structuré
2)	Polyhex graph Hexagonal animal }	Benzenoid system
3)	Transfer matrix, Local states [13]	Kekulé structures
4)	Conjugatead circuit [21]	A cycle of alternating single and double bonds.

5)	Permutation integral [12]	Interaction between two Kekulé structures.
6)	Factor graph [14]	Kekulé structure-model
7)	Maximal independent set of vertices [22]	Clar structure
8)	Nonadjacent hexagons	Resonant hexagons.
9)	Caterpillar tree, $T_{}$	Unbranched benzenoid
10)	Young diagram	
11)	Line graph of T [17]	Clar graph; Branched and unbranched systems
12)	Supermatch [37]	Supersextet.
13)	Noninteracting perfect matchings [38]	Localized double bonds
14)	King polyomino [19]	Benzenoid system containing no more than 4 hexagons
15)	σ (B; 1) [39]	K = number of Kekulé structures.
16)	ξ (B; 1) [40]	ζ = number of Clar structures.
17)	Inner dual, dualist [2]	Annelation of hexagons in a benzenoid hydrocarbon.
18)	Isoarithmic polyhex graph [41]	Benzenoid systems possessing the same value of K.
19)	Cospectral caterpillar [42]	Unbranched benzenoid hydrocarbons with the same K.
20)	The John-Sachs determinant [43]	K
21)	Fibonacci numbers [44]	K values of the zigzag benzenoid hydrocarbons: benzene, naphthalene, phenanthrene, chrysene, ...
22)	Bipartite graphs (the dancing problem [45] in graph theory).	Alternant hydrocarbons.
23)	Pairing theorem [46]	Distribution of energy levels in benzenoid systems.

Some of the above terms have already been mentioned others can be guessed or found in the literature. The purpose of Table 5 is to show "at a glance" the richness of the "benzenoid theory", in novel as well as classical concepts and to suggest in a number of cases

where it is apparent, that some of the "key" words occur in other chemical disciplines, outside benzenoids. Hence, while a concern with benzenoids to some may appear as an "aberration", already the list of key words illustrates that the "deeper" one goes into the subject the "broader" the subject becomes, something typical of many other scientific disciplines. But there is one additional factor worth pointing out which to some extent justifies initial focusing of graph theorists on hydrocarbons. Surely, chemistry involves much more than carbons and hydrogens. If one could extend such considerations by including also oxygen and nitrogen much of the interest of organic chemistry could be considered. However, hydrocarbons represent the "testing" ground in chemical graph theory for many algorithms constructions and designs. If the algorithm fails in case of hydrocarbons, whether this involves isomorphism, automorphism, enumeration, special construction, "rational" design of invariants, etc., there is clearly no point in considering its generalizations to heterographs (i.e., colored graphs). Hence continuing interest in benzenoids, and in hydrocarbons in general, is legitimate and is a promising springboard for extending graph theory to heterocyclic chemistry, inorganic chemistry, and even biochemistry and biophysics. Many such problems would require extension of the concept of graphs to "absorb" information on 3-dimensionality of molecules, and some initial results in that direction have been recently reported by one of the authors [47], but again only for hydrocarbons!

2.1.2 Counting polynomials

A convenient way of expressing the outcome of a particular enumeration is through a counting polynomial. Table 6 portrays some of the more important polynomials adopted and defined for graph-

Table 6: Counting polynomials of the paths L_2-L_7, Independence polynomials of L_1-L_6 and Sextet polynomials of benzene, naphthalene, chrysene, picene and fulminene.

$1 + x$
$1 + 2x$
$1 + 3x + x^2$
$1 + 4x + 3x^2$
$1 + 5x + 6x^2 + x^3$
$1 + 6x + 10x^2 + 4x^3$

theoretical problems. In Fig. 7 we show four types of graphs: a caterpillar, T, a Clar graph Λ, a benzenoid graph, B and a polymino, P,

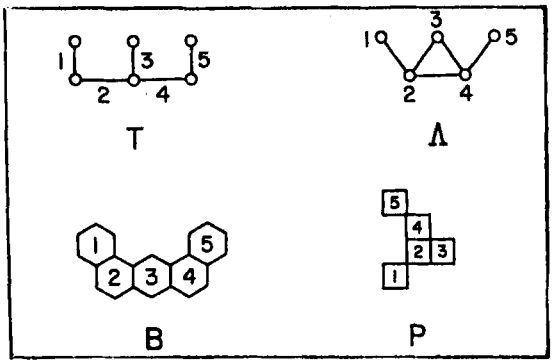


Fig. 7: Four equivalent graphs, arbitrarily labeled. There is a one-to-one correspondence between edges in T, vertices of Λ, hexagons of B and the squares making board P.

Oberve that the following "adjacency" matrix relates equally to the four distinct objects of Fig. 7.

$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Labels 1-5, which are in fact arbitrary labels for edges, vertices, hexagons and cells in each of the cases illustrated. If we now take the above matrix to represent the adjacency for each of the four situations, then we have induced very specific adjacency relationships. We obtain respectively:

Graph	Adjacency
Caterpillars	edge incidence
Line graph	vertex adjacency (usual adjacency matrix).
Polyhex	ring "resonance" (to be explained soon).
Cell graph	"non-attacking" kings of chess game (to be explained).

Hence, in summary, different objects and distinctive definitions of adjacency, can be represented by the same adjacency matrix !

Resonant sextets

The concept of "resonant rings" follows from Clar's representations of benzenoids [48] in which one identifies "pi-sextets", that coincide with six pi-electrons of various individual benzene rings in polycyclic-benzenoids. If a pi-sextet is represented by an inscribed circle we can immediately see that, for example, rings 1 and 4 in the benzenoid considered are "resonant", i.e., both rings can simultaneously be "sextet" rings. However, rings 2 and 4 are not resonant, because we cannot

complete the resulting valence structure by "coupling" all the remaining carbon centers into a Kekulé valence structure. Hence the elements of $\underline{A}$ describe "resonance relations" of hexagons in a benzenoid system as follows

$$a_{ij} = \begin{cases} 1 & \text{if hexagons i, j are nonresonant} \\ 0 & \text{otherwise} \end{cases}$$

The above relations define a graph, which is a "line graph" of the associated caterpillar [17]. Line graph L(G) is defined by relations in which edges of G represent vertices in L(G) and adjacency in L(G) corresponds to includence of edges in G. Such graphs can be constructed for all types of benzenoids, i.e., contrary to caterpillars they are not restricted to unbranched systems.

Polyomino graph

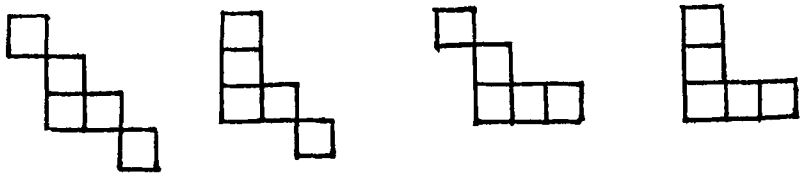
A polyomino or a "square" animal is composed of square cells of the same type. From the tables of the king and domino polynomials published by Motoyama and Hosoya [19] we see that there are 35 hexaminos and 86 heptaminos (i.e., polyominos with 6 and 7 cells respectively). The adjacency matrix A describes how two or more king pieces of chess game may occupy cells of more generalized boards in which we combine polyominos such that they can share vertices not only edges, Thus:

$$a_{ij} = \begin{cases} 1 & \text{if cells i and j } \underline{\text{cannot}} \text{ have simultaneously two "nontaking" kings} \\ 0 & \text{otherwise} \end{cases}$$

Since the rules of chess stipulate that two nontaking kings cannot be in adjacent "squares" the above definition can be written simply as

$$a_{ij} = \begin{cases} 1 & \text{if cells i and j are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

Two cells on a square are said to be adjacent it they share at least one corner. Observe that the above definition of a polyomino does not determine the shape of the board, but just the adjacency. Hence in our example (Fig. 7), P can be pictorially represented by any of the following shapes:



Non-uniquness of representing A is also reflected in the equivalence between pairs of benzenoid hydrocarbons such as:

Both of which are associated with the same sextet adjacency matrix. The objective of this part of the review is to present a topological theory which unifies many "seemingly" unrelated graphs and polynomials of chemistry and physics. This will be done by specifying equivalence relations with which we can associate equivalent objects, graphs and "graphoids" (Such as polyominos). In particular we will seek recursive relations between members of different size and will illustrate how equivalent recursions valid for one group of mathematical objects or graphs can be used for another set of objects related by equivalence. We will consider also structures with "supersextets" [37] and implication of "super" delocalizability. It will be conjectured that, in principle,

larger graphs can be studied in terms of smaller equivalent graphs (objects). Such equivalence, naturally does not extend over all properties of such graphs, hence we have to investigate which properties of interest can in this way be studied. Such a direction of investigation can be legitimately viewed as a novel approach to data reduction, since the equivalent objects which form the basis of our study of large graphs will, as a rule, be defined by fewer input data. "Classical" data reduction relates to extraction of common properties from a larger collection of data: The novelty considered here consists in data reduction obtained by extraction of the same data of interest from simpler objects. We end this section by a statement of Max Planck which illustrates the importance of unifying seemingly different phenomena: "The chief problem in every science is that of endeavouring to arrange and collate the numerous individual observations and details which present themselves, in order that they may become part of one comprehensive picture".

2-1-3 Equivalence relations

Let us consider more closely the set of objects of Fig. 7 associated with the same adjacency matrix $\underline{A}$. We will refer to these as the set $\{ T, \Lambda, B, P \}$ standing, respectively, for a caterpillar tree, a Clar graph ($\Lambda = L(T)$) a benzenoid graph and a king polyomino. The graph invariants that we will consider in each case are as follows:

$$\text{a set of edges in T} : \{E(T)\} = \{ E \} = \{ e_i \mid i = [1,5] \}$$
$$\text{a set of verticies in}_\Lambda : \{ V(\Lambda) \} = \{ V \} = \{ v_i \mid i = [1,5] \}$$
$$\text{a set of hexagons in B:} \{ H(B) \} = \{ H \} = \{ h_i \mid i = [1,5] \}$$
$$\text{a set of cells in P} : \{ C(P) \} = \{ C \} = \{ c_i \mid i = [1,5] \}$$

With each of the above we can associate the same polynomial, viz.

$$1 + 5\,x + 5\,x^2 + x^3$$

which enumerates in the caterpillar tree the number of disjoint edges, the number of independent vertices in the line graph, the number of resonant sextets in the benzenoid hydrocarbon, and finally the number of nontaking kings in the polyomino. In Fig. 8 we illustrate on a smaller graph of benzanthracene the equivalence of the count of edges, vertices, resonant hexagons and "nontaking" placements of king pieces on cells in T, Λ, B and P respectively. By convention the first (constant) term is equal to 1. The coefficient of the linear term counts the number of edges, vertices, hexagons and cells respectively in T, Λ, B and P graphs. The coefficient of the quadratic term counts the number of pairs of nonadjacent edges, the number of pairs of nonadjacent (so called independent) vertices, the number of resonant pairs of sextets and finally the number of two nontaking kings. Higher

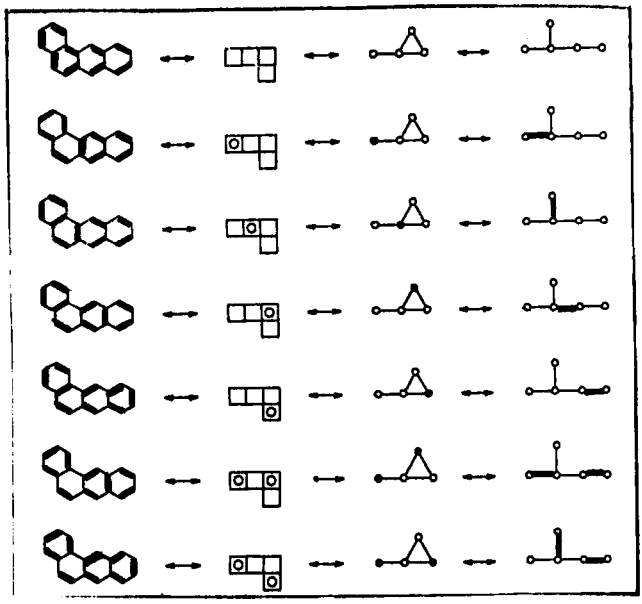


Fig. 8: Equivalence of counting. The polynomial $1 + 4x + 2x^2$ describes any of the four objects.

terms when present can similarly be interpreted. Hence the four apparently different counting polynomials are indeed descriptions of similar combinatorics.

The above illustrates the kind of equivalence relation we are interested in. Then, for the graphs of Fig. 7 we have the following identity:

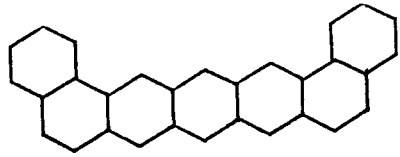
$$H(T; x) = \omega(\Lambda; x) = \sigma(B; x) = K(P; x) = 1+5x + 5x^2 + x^3$$

Although it appears, now, that the above distinct topics have a common enumerative basis it should be realized that most of the known results, over two decades at least, were obtained independently for each of the above topics, which were only recently recognized as equivalent and thus offering a unified picture of diverse physicochemical explorations. In order to proceed with proofs for the above intuitive notions on equivalence we need the following:

Lemma: Bijective (one-to-one, onto) functions [15] exist between each of the above mentioned sets of graph invariants; E, V, H, C. In other words we can write the relation as follows:

$$R = \{(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4), (X_5, Y_5)\}$$

where X, Y stand for any of E, V, H, or C. Observe that we have already selected labeling of graph invariants so that there is a one-to-one correspondence between the invariants that make equivalence (in this relatively simple illustrative case) rather apparent. But, in general, labels need not be available and verification that two objects, such as T, Λ, B and P, are equivalent requires a careful analysis. For example, we may be presented with a benzenoid, such as



and may have first to identify the corresponding board, if it exists (see later). From the illustrations used to introduce these various mathematical objects it is not apparent that there are some restrictions imposed on the equivalence. To have a counting polynomial for resonant sextets implies $K \neq 0$, i.e., the existence of Kekulé structures. Hence polyhexes considered must be benzenoid (according to our definition of benzenenoids). However not all benzenoids lead to caterpillars: restriction here is not only that a benzenoid must be catacondensed but also unbranched. Furthermore restriction in the number of linearly fused rings (as will be elaborated later) is important. With any caterpillar we can associate a benzenoid hydrocarbon but the reverse is not the case. The restriction on boards is that they must be superimposable on a regular grid, such as a chess board, which restrict the number of "adjacent" cells. The maximal arrangement of cells then corresponds to tetracene. Fig. 9 illustrates restriction to the "size" of the benzenoid system.

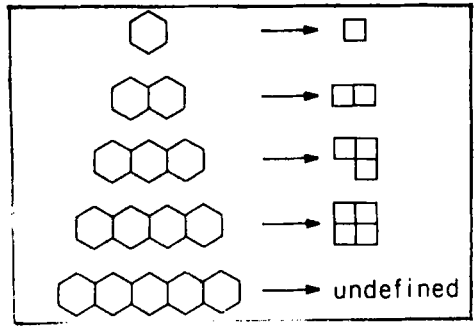


Fig. 9: Illustration of the restriction on equivalence as a function of the number of linearly fused hexagons.

There are no restrictions on construction of line graphs, but of course, not every line graph is of interest here.

In establishing equivalence between qualified invariants in different objects, such as {E} in caterpillars, and {V} in line graphs it is arbitrary to choose which set is the domain and which is the range, because of the existence of one-to-one correspondence. If {V} is taken to be the image of {E}, then a function $f_{E,V}$ is defined for

$$f_{E,V} : \{E\} \longrightarrow \{V\}$$

Further, if $\underline{A}$ is still the incidence matrix of T and the adjacency matrix of Λ, then f(E,V) is a function which transforms a given caterpillar into the associated line graph [5], i.e.

$$f(E,V) : T \longrightarrow \Lambda \text{ or } L(T) = \Lambda$$

The above means that f transforms every edge in T into a unique vertex in L(T), i.e.

$$f(E,V): T \longrightarrow \Lambda \longrightarrow \{f(e_i, v_i): e_i \longrightarrow v_i \quad i = [1,5]\}$$

Analogous mappings can be envisaged for {E} $\longleftrightarrow$ {H}; {E} $\longleftrightarrow$ {C}; {V} $\longleftrightarrow$ {H}; {V} $\longleftrightarrow$ {C}; {H} $\longleftrightarrow$ {C}; where the double headed arrow means that both f and f^{-1} can be defined provided one conforms to restrictions that stipulate when alternative mathematical objects are defined for the benzenoid considered.

We should add, that while we confine our discussion to the four types of graphs: caterpillars, line graphs, selected benzenoids, and qualified boards, the analysis holds generally for any other objects for which analogus "adjacency" relations can be defined. For example consider in Fig. 10 the caterpillar T and the associated "walk around" code of

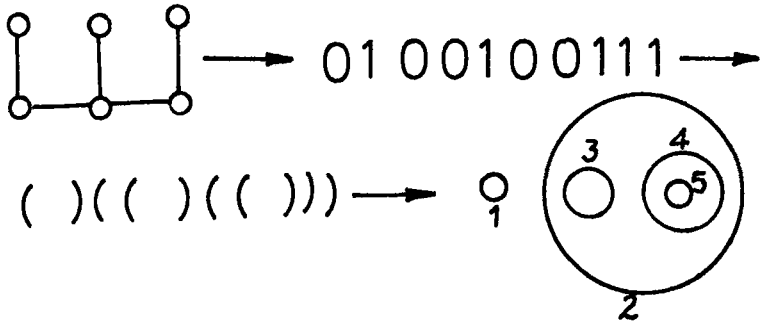


Fig. 10: Walk codes and equivalent contours possessing the same incidence matrix $\underline{A}$ of the indicated comb caterpillar.

Read [49], in which one assigns 0 or 1 to each edge as one walks around the graph starting with edge 1 and moving clockwise.

The initial 01 code means that edge 1 was "visited" (or "looked at") from both sides before any other edge was visited. We continue with 001, signifying the first encounter of edge 2 (first digit 0), and "visiting" edge 3 from both sides (subsequent 01). By completing the "walk around" one obtains the code

0 1 0 0 1 0 0 1 1 1

Clearly the form of the code depends on the starting point of the "walk" and the direction of circling the tree. The above code can be "translated" into a set of inscribed circles, by first replacing 0, 1 labels by left and right brackets leading to "nested loops" specified as

( ) (( ) (( )))

We can now speak of "adjacent" loops and define the adjacency via:

$$a_{ij} = \begin{cases} 1 & \text{if the "loops" are sequential} \\ 0 & \text{otherwise} \end{cases}$$

By "sequential" we mean that one loop is executed following immediately another, or was executed immediately preceeding another, in both cases such events are taken as "adjacent".

If the above brackets are "closed" by connecting adjacent left and right bracket first and then continuing the process by ignoring the brackets already connected we obtain a geometrical diagram representing a set of inscribed nonoverlapping circles. This particular correspondence between trees and nonintersecting contours, of interest in the field of partial differential equations, has found some applications in chemistry too! However in contrast to the standard association of trees and such contours by associating with various "regions" the individual vertices we will depart and assign labels to the individual contours, not regions. The labels associated are choosen to correspond to the edges of the initial caterpillar. This leads to yet another illustration of mathematical objects assocaited with the same adjacency materix $\underline{A}$ if we define:

$$a_{ij} = \begin{cases} 1 & \text{if contours are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

"Adjacent" here means that no other contour is in between i and j.

Proposition 1

A "restricted" equivalence relation can be defined and physically realized on the set M;

$$M = \{ T, \Lambda, B, P \}$$

Proof:

In purely mathematical terms an equivalence can be defined on an arbitrarily chosen subset of M. Thus, for example, the relation:

$$R = \{ (T, T), (\Lambda, \Lambda), (B, B), (T, \Lambda), (\Lambda, T), (\Lambda, B), (B, \Lambda), (T, B), (B, T) \}$$

is an equivalence relation on $\{T, \Lambda, B\}$. One can realize the physical meaning of R by consideration of the lemma previously stated. The reflexive part of R is obvious since every graph is related to itself. The fact that there is a bijection (one-to-one, onto) of $E \longleftrightarrow V$ etc. induces the symmetric condition of R. Finally, it follows naturally from the definition of the polynomials, $\{ H(T; x); \omega(\Lambda; x), \sigma(B; x), K(P; x) \}$ and the corresponding "adjacency" relations (c.f. matrix A) of their associated graphs that the existence of:

$$f(V, T): \Lambda \longrightarrow T \quad \text{and} \quad f(E, H): T \longrightarrow B$$

leads to

$$f(V, H): \Lambda \longrightarrow B$$

which establishes the transitive character of R. Indeed the mapping $f(V, H): \Lambda \longrightarrow B$ is implied in the definition of Clar graph [4] while the mapping $f(V, T): \Lambda \longrightarrow T$ is a function that inverts a line graph back into the original graph and $f(E,H): T \longrightarrow B$ maps a set of incident edges into a set of nonresonant hexagons and a set of nonincident edges into a set of resonant hexagons.

Analogous R relations exist for all possible subsets of M as well as for the augmented set S in which "nested loops" and "inscribed contours" and other such possible objects may be possible.

2.1.4 Equivalence classes

An equivalence relation such as R defined above partitions a set of graphs into disjoint subsets called equivalence classes; the elements of each equivalence class are related. The concept is illustrated on a set of graphs shown in Fig. 11. The equivalence relation is seen to partition the set of twelve graphs of Fig. 11 into three blocks as shown in Fig. 12. Every block is characterized by one polynomial.

For example, for the middle block one has

$$\sigma(b_1; x) = K(p_1; x) = \omega(\lambda_1; x) = H(t_1; x) = 1 + 4x + 2x^2$$

Therefore $\sigma(b_1; 1) = 7$, i.e., there are seven perfect matchings of b_1 (i.e. seven valence-bond Kekulé structures), there are seven ways of arranging nontaking king on p_1 (not exceeding two kings), there are

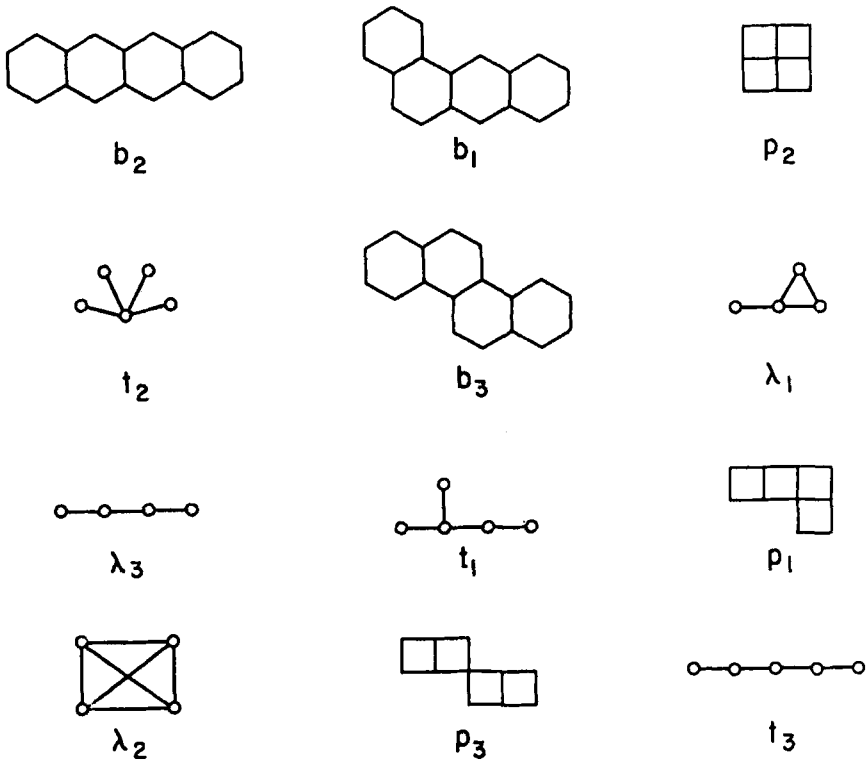


Fig. 11: A set of twelve graphs comprising four types: a caterpillar, a polyhex graph, a Clar graph and a polyomino graph.

seven matchings in t_1 (not exceeding a maximum of 2-matchings), finally there are seven independent sets of vertices in λ_1. Thus the combinatorics of one element in a subset is shown to be preserved throughout the rest of the elements.

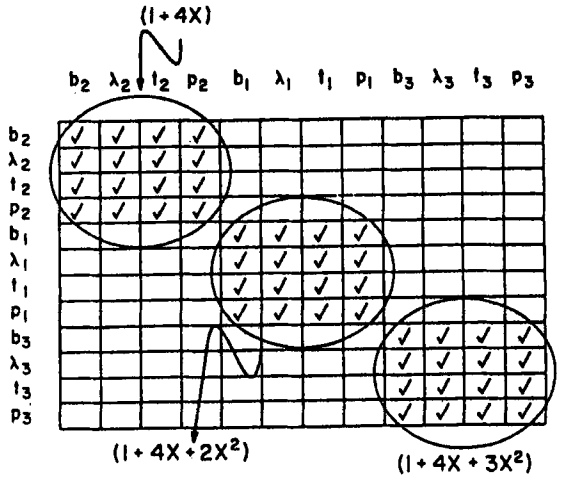


Fig. 12: The three equivalence classes of the set of graphs shown in Fig. 11. The equivalence polynomials are written in each block.

2-1-5 Equivalent graphs and equivalent polynomials

Subsets of graphs which constitute equivalence classes will be called equivalent graphs and their polynomials, equivalent polynomials. It takes a little thought to realize that an equivalence relation (with transitive closure [50] on a set of n equivalent graphs requires the existence of n! mappings of selected graph invariants onto each other. Thus for the set S = { T, Λ, B, P} one must identify 24 (=4!) mappings, three of which are

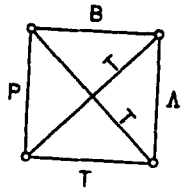
$$T \longrightarrow \Lambda \longrightarrow B \longrightarrow P$$

$$T \longrightarrow B \longrightarrow \Lambda \longrightarrow P$$

$$T \longrightarrow P \longrightarrow \Lambda \longrightarrow B$$

where in general $T \longrightarrow \Lambda \leftarrow \{E\} \longrightarrow \{V\}$ i.e. that the edges of T are in exact one-to-one correspondence with the vertices of Λ (c.f. Lemma) and so on for other mappings. In other words each graph is viewed as a subset of elements (selected graph invariants) such that mapping of the whole graph into another graph. It must be emphasized that equivalent graphs keep adjacency relations between selected invariants as explained above.

A pictorial illustration of equivalent objects is depicted below where the individual elements occupy the edges of a complete graph.



Such a representation is particularly useful when one recalls that there is a natural equivalence relation on the set of edges contained in a cycle [51].

Proposition 2

Define a general polynomial function F(G;x) Let S be a set of equivalent graphs { G_1, G_2 , ... } . Then $F(G_1; x) = F(G_2; x) = ...$ if $G_i \in S$.

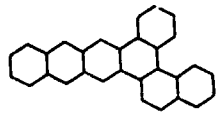
Proof:

If G_1, G_2, ..., are equivalent graphs there must be, by definition, one-to-one onto mappings defined for selected sets of their graph invariants. Therefore all such graphs have identical numbers of graph invariants [52]. Furthermore in each case the set of invariants have "equivalent" adjacency relations, e.g. e_1 is related to e_5 (in T) in the same way as v_1 is related to v_5 (in Λ). Then $\theta(G,k)$'s enumerate identical combinations of different objects.

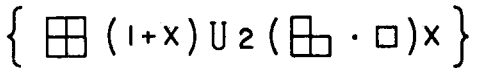
The existence of cospectral gaphs [53] prevents one from an "only if" statement in proposition 2. Therefore if $F(G_1; x) = F(G_2; x)$ one concludes that either G_1 and G_2 are equivalent graphs (i.e. { G_1,G_2} is an equivalance class or G_1 and G_2 are cospectral graphs [53].

An observation of mapping a polyhex graph into a (king) polyomino graph:

As illustrated earlier it is possible to find an equivalent king polyomino to linear chains of hexagons of only up to 4 hexagons (c.f. Fig. 9). Also, it is not possible to define a king polyomino for a branched or a pericondensed benzenoid hydrocarbon. Nevertheless with the use of the appropriate recursive relations and necessary transformations [54] one can describe a particular graph, usually of large size, in terms of a set of smaller equivalent graphs and thus offer an easiler route to understanding the combinatorics of the larger system. For example it can be shown that the polyhex shown below:



is indeed equivalent to the set of pseudo- [55] king polyominos:



in fact for the above set K(P; x) is found (by inspection) to be:

$$(1+4x)\,(1+x) + 2(1+3x)\,(1+x)x = 1+7x+12x^2+6x^3$$

which is indeed the sextet polynomial, $\sigma(B_7; x)$.

2-1-6 Equivalent recursive relations, graph-inter relations

It is not difficult to make use of known recursive relations [54] on the members of M (or other sets of equivalent graphs) to write identities in their graphic notations where the graph stands for its coresponding polynomial, e.g. T = H(T; x), B = (B; x) etc.

One observes the following interesting relations:

(i), (ii) ⟶ $(t_1 R \lambda_1)$; $(t_1' R \lambda_1)$ where in general (aRb) means that a is related to b. Actually $L(t_1) = \lambda_1$; $L^{-1} \lambda_1 = t_1$; $L(t_1') = \lambda_1'$; $L^{-1} \lambda_1' = t_1'$ where L(G) is the line graph of G.

(i), (iii) ⟶ $(t_1 R b_1)$, $(t_1' R b_1')$

in fact, $H(t_1; x) = \sigma(b_1; x)$;

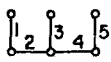
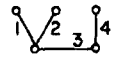
$H(t_1'; x) = (b_1'; x)$

(i), (iv) ⟶ $(t_1 R p_1)$; $(t_1' R p_1')$

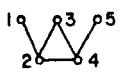
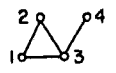
Actually

$H(t_1; x) = K(p_1; x)$

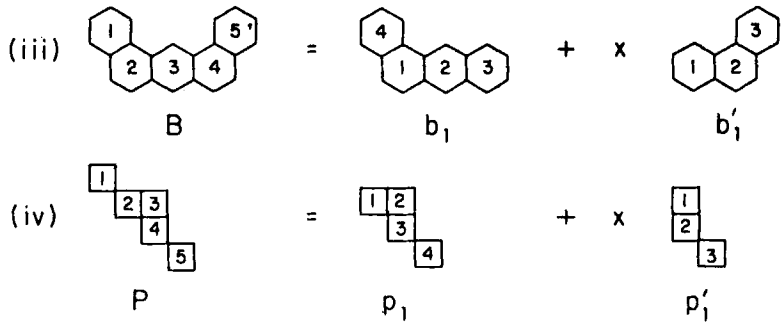
$H(t_1'; x) = K(p_1'; x)$

(i)  =  + x 

T t_1 t_1'

(ii)  =  + x 

Λ λ_1 λ_1'



Three other analogous relations might be written for the above set of recursive relations. Using similar reasoning to that used of Lemma and Proposition 1 it is not difficult to demonstrate that both $(t_1, \lambda_1, b_1, \mu_1)$ and $(t_1', \lambda'_1, b_1', p_1')$ are sets of equivalent graphs. This reasoning leads to:

Proposition 3

Recursive relations can be written for a set of equivalent graphs which generate different sets of equivalent graphs. Such recursive relations which lead to sets of equivalent graphs might be called equivalent recursions. It may be instructive to speak of a set of equivalent garphs as contained in another set (of the same cardinality) of equivalent garphs (of larger sizes). Thus one may write

$$\{t_1, \lambda_1, b_1, p_1\} \subset \{T, \Lambda, B, P\}$$

whence $\{t_1, \lambda_1, b_1, p_1\} = \{\{e_1\}, \{v_1\}, \{h_1\}, \{c_1\}\}$

i.e. the graph notation stands for the subset of its selected invariants. For example $\{e_1\}$ is the subset of edges in t_1 and so on. Analogously $T = \{E\}$; $\Lambda = \{V\}$; $B = \{H\}$; $P = \{C\}$.

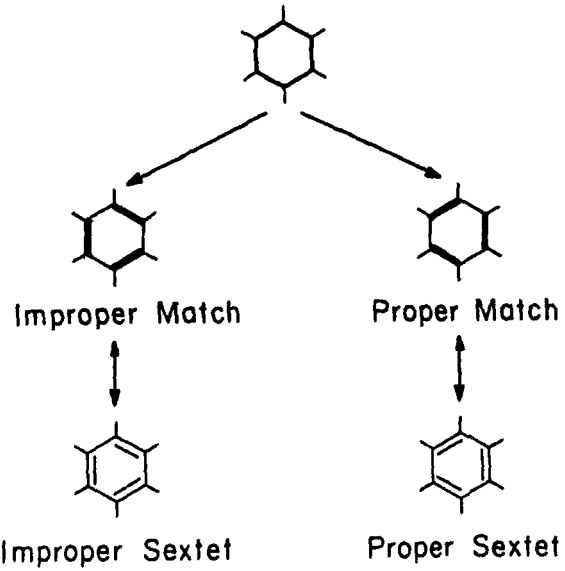
Similarly one can write

$$\{t_1', \lambda_1', b_1', p_1'\} \subset \{T, \Lambda, B, P\}$$

where $\{E\} = \{e_1\} \cup \{e_1'\}$ which in graph notation becomes $\{T\} = \{t_1\} \cup \{t_1'\}$. Analogous equations may be written for $\{\Lambda\}$ $\{B\}$ and $\{P\}$.

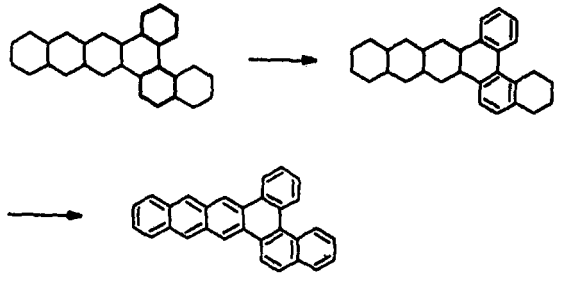
On the Topological Theory of Polyhex Graphs
"Super-Sextets" and "hypo-Sextets":

In their analysis of the topological dependency of the aromatic sextet in polycyclic benzenoid hydrocarbons Ohkami and Hosoya [37] arbitrarily define two types of ring perfect matchings viz., proper and improper as shown below:



Every "line" in the perfect matching corresponds to two pi-electrons, whence three lines correspond to an (aromatic) sextet of electrons.

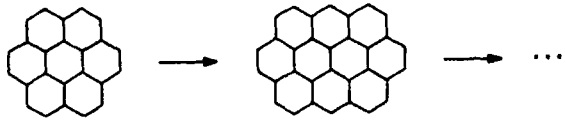
When hexagons with bold edges are transformed into proper sextets (= proper matchings) and the rest of the edges are matched so that no other hexagon contains a proper sextet, a Kekulé structure results. For example:



There seems, then, to be a one-to-one correspondence between the sextet patterns and the Kekulé patterns (a Kekulé structure is nothing else but a perfect matching [56]). This is not true, however, since such a one-to-one correspondence is lost in the following major classes of polyhexes:

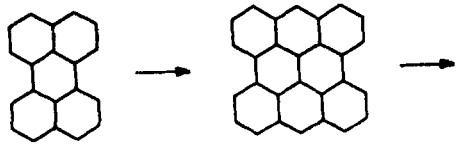
Class A:

Coronene and higher members:

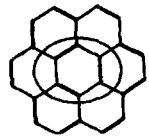


Class B:

Perylene and higher analogues:



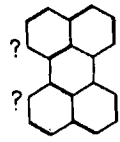
In class A one (or more) Kekulé (perfect matching) pattern is lost because of the existence of what Ohkami and Hosoya [37] call a "super ring" or a "super-sextet" which may be called a "super-matching" as illustrated below:



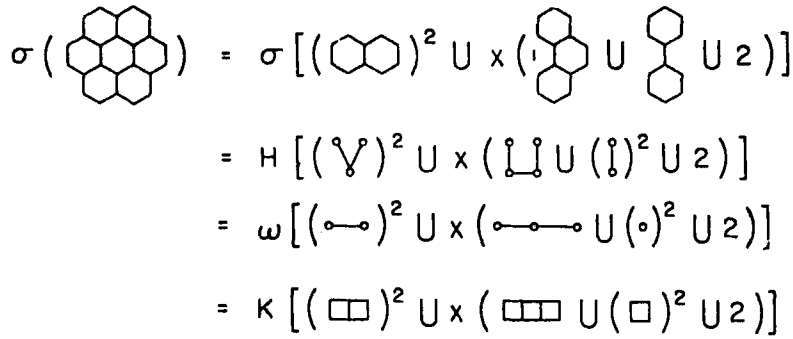
a super sextet = a super match

In fact, σ(coronene; x) = 1 + 8x + $9x^2$ + $2x^3$; the 8x suggests an "extra" ring.

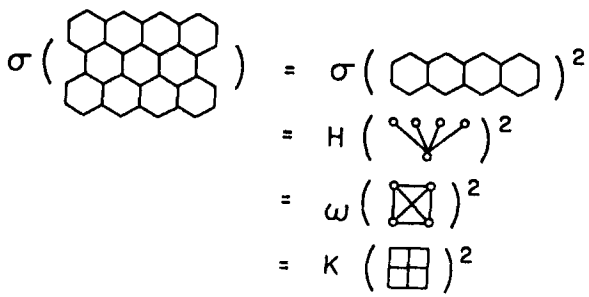
In Class B the middle row of the hexagons cannot have ring sextets, viz.,



Acually σ (perylene; x) = 1 + 4x + $4x^2$. The term 4x suggests the presence of 4(rather than 5) rings. This class might therefore be described as having a "hyposextet" (or a "hypo ring"). This topological dependency of the perfect match pattern is best understood in terms of recursive relations [54]. It can be shown that the sextet polynomials of polyhexes which have super rings are best expressed in terms of polynomials of equivalent graphs. Thus for example:



Polyhexes with "hypo rings" can also be understood in terms of their equivalent graphs. For example

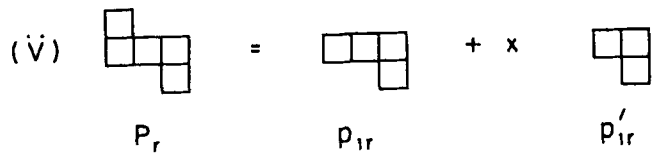


On Rook Boards [57]

A "rook" is a chessboard piece which "takes another piece on the same row (or column)". The enumeration of the number of ways of placing k nontaking rooks on a chessboard, Pr is useful in counting the number of permutations of objects when there are restrictions on the positions they can occupy. Results from Rook theory were found to be useful in the topological theory of resonance [57] in connection with the matching polynomials. In a fairly recent development Godsil and Gutman [57] demonstrated that there is a one-to-one mapping between a labeled bipartile graph $G = G(P_r)$ with $a + b$ vertices and boards with a rows and b columns such that a cell $c_{ij} \in P_r$ iff vertex $v_i \in \{a\}$ is connected to vertex $\omega_j \in \{b\}$ in $G(P_r)$. Since the tree T is bipartite the equivalent rook board of it exists and is drawn below

P_r

There are five ways of placing two nontaking rooks on $P_r(T)$, viz., (13), (14), (15), (25) and (35) and only one way of placing three such rooks; (135). The numbers assigned to the individual cells correspond to the edges of T as labeled, so the mapping in this case is also bijective. An equivalent recursion might then be added to the set (i) - (iv) above, viz.,



2-1-7. On young diagrams [30,31], general ordering [58] of equivalent graphs:

Ruch and Schönhofer [31] defined a (partial ordering) relation among a set of Young diagrams correlating to the set of irreducible representations of the symmetric group. Such a partial ordering relation is best understood in terms of ordering rules of Muirhead [58]. Thus, if one denotes the number of cells (squares) in successive rows of a Young diagram Y_n by $a_1, a_2, \cdots, a_n$ and the corresponding numbers in Y'_n by $a'_1, a'_2, \cdots, a'_n$ then Y_n is greater than or equal to (and is comparable with) Y'_n if

$$a_1 \geq a'_1$$

$$a_1 + a_2 \geq a'_1 + a'_2$$

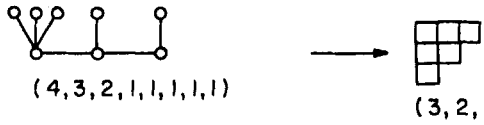
$$\vdots$$

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a'_i$$

In the mean time Gutman and Randić [10] defined an algebraic charcterization of skeletal branches of trees. They listed the degrees of vertices of a tree in descending order, $d_1, d_2, \ldots, d_i$ where $d_1 \geq d_2 \geq \cdots \geq d_i$. Using similar set of eqns. they [10] were able to order sets of trees containing 8, 9 and 10 vertices. The ordering of a set of Young diagrams containing n squares can be made to exactly match the hierarchy of ordering a set of (n+2) trees if: [10]

(a) Information on terminal vertices is suppressed.
(b) The remaining vertex degrees are reduced by one.

For illustration we consider a tree and the corresponding Young diagram:



where the numbers in the left parentheses are vertex-degrees listed in descending order while those of the right parentheses are the number of cells in rows 1, 2 and 3 of the Young diagram corresponding to the tree shown. One observes that for every tree there is <u>one</u> Young diagram, however the tree corresponding to a given Young diagram is <u>not</u> unique. Thus, there are, e.g., three trees coresponding to Y(3,2,1) i.e. the Young diagram shown above, viz.

T (3,0,2) T (3,1,1,) T (2,2,1)

The numbers in parentheses are the numbers of monovalent vertices in T counted from left to right)

All the above trees have vertex degrees of (4,3,2,1,1,1,1,1)• One might speak of a set of trees belonging to a given Young diagram, thus one may write:

$$\{T_3(3,0,2)\cdot\ T_3(3,1,1),\ T_3(2,2,1)\} \approx Y(3,2,1)$$

Similarly

$$\{(T_4(2,0,0,2), T_4(2,0,1,1),\ T_4(2,1,0,1),\ T_4(1,1,1,1)\} \approx Y(2,2,1,1)\cdot$$

Other such expressions might be obtained by investigating other sets of trees•

The observation that there is one-to-one correspondence between the ordering of a set of Young diagrams containing n squares and a set of trees containing (n+2) vertices suggests the following:

Conjecture 1:

A greater relation can be defined between sets of equivalence classes of graphs• Each equivalence class is described by a unique Young diagram as dictated by the vertex structure of the tree of the given class•

Thus the hierarchical ordering diagrams described in Ruch's paper [31] can be made to order equivalence classes of graphs• One, then, anticipates that the physical properties of the molecules (benzenoid hdyrocarbons for polyhex graphs alkanes for trees etc•) can also be ordered and predicted• Furthermore, one expects that every given class of equivalent graphs ("equivalent molecules") to have very similar (if not indeed identical) "global" properties•

Proposition 4 [5]

Polyhex graphs in which no vertex is common to more than two hexagons are equivalent to caterpillar trees and vice-versa•

Proof

Let B be a polyhex graph in which no vertex is common to more than two hexagons and T be its equivalent tree. Then by definition,

$$\sigma(B;x) = H(T;x)$$

Then the resonance-relations of the hexagons $\in B$ exactly match the edge-incidence relations of the-edges $\in T$. Since a caterpillar tree $T_n(m_1, m_2, \cdots, m_n)$ is constructed by the addition of m_i monovalent vertices to v_i of a path P_n the proposition is proved, i.e. $T_n(m_1, m_2, \cdots, m_n)$ corresponds to a polyhex in which there are $m_1, m_2, \cdots, m_n$ hexagons in linear annelation (m linearly annelated hexagons are essentially nonresonant).

The fact that several trees may belong to the same Young diagram generates "degenerate" equivalence classes. The individual classes are associated with the equivalent polynomials ($\equiv H(T;x) = \omega(\Lambda;x) = \sigma(B;x) = K(P;x) = R(P_r;x)$). The limits of the hierarchy are classes corresponding to the linear acenes, star trees and perfect graphs; and the class corresponding to the zigzag type polyacenes both the caterpillars and Clar graphs of which are simple paths. The king boards of this class exist as strings of linearly annelated squares. This interesting equivalence class generates most of the Fibonacci-number-relations in chemistry [7,44]. Fig. 13 shows the general form of the resulting hierarchy.

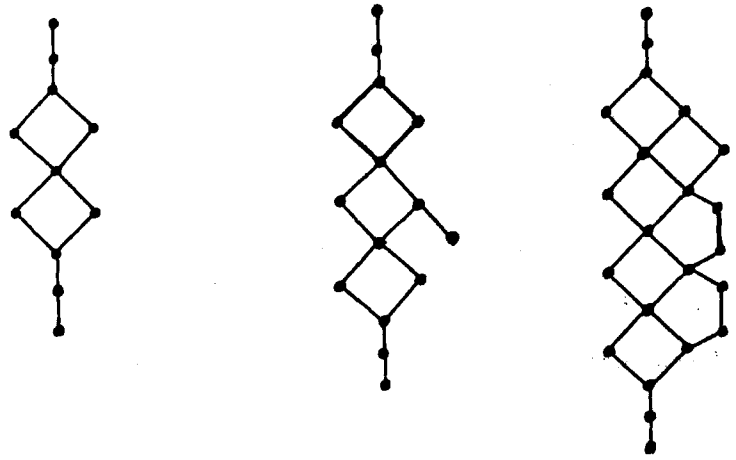


Fig. 13: General hierarchies based on partitions of 6,7 and 8 respectively. Top levels correspond to linear acenes, star trees and perfect graphs while bottom levels are those of the zigzag polyacenes and paths.

2.2 HOMOLOGOUS SERIES OF GRAPHS

A set of graphs of the same type in which the number of vertices increases in a regular way defines a homologous set of graphs. For example, the linear acenes viz., benzene, naphthalene, anthracene, etc. constitute a homologous set of graphs. It can be easily shown that there is an equivalence relation between the members of such sets. Sometimes the graph-theoretical properties of a member of a homologous series can always be predicted if those of the two immediately preceeding members are known. I.e., the graph-theoretical properties of a homologous set can be described by a recursive relation. For example the number of Kekulé structures of a member of the zigzag polyacene containing j hexagons recur in the following way:

$$K(B_{j+2}) = K(B_{j+1}) + K(B_j); \quad j > 1$$

where $$K(B_j) = F_{j+1} = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{j+2} - \left(\frac{1-\sqrt{5}}{2}\right)^{j+2}\right];$$

One recalls the above equation as the Binet formula of Fibonacci numbers [44]. Furthermore, for such a member of homologous set, the number of selections of k resonant hexagons is given by [34,44]

$$r(B;k) = \binom{j+1-k}{k}$$

A more well known result is the case when a homologous set of linear acenes is defined where $K(B_j) = j+1$. In Fig. 14 we show a number of homologous sereis of benzenoids. Table 6 lists Clar [27] and sextet [39] polynomials for such equivalence classes.

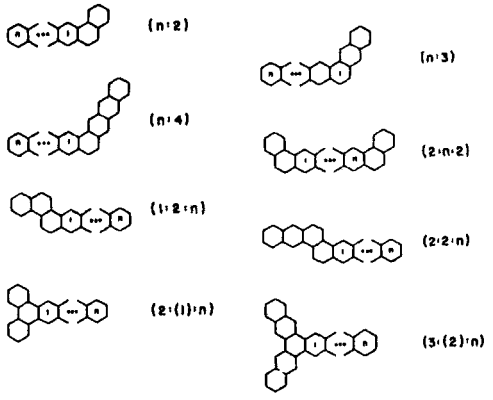


Fig. 14: Homologous series of benzenoid systems.

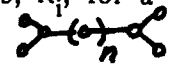
Table 6: Clar and sextet polynomials of equivalent graphs shown in Fig. 14.

Class	Clar polynomial	Sextet polynomial
	$CR(G;x)$	$\sigma(G;x)$
$(n:2)$	$x+nx^2$	$1+(n+2)x+nx^2$
$(n:3)$	$x+2nx^2$	$1+(n+3)x+2nx^2$
$(n:4)$	$x+3nx^2$	$1+(n+4)x+3nx^2$
$(2:n:2)$	$2x^2+nx^3$	$1+(4+n)x+(2n+3)x^2+nx^3$
$(1:2:n)$	$(2n+1)x^2$	$1+(n+3)x+(2n+1)x^2$
$(2:2:n)$	$2x^2+3nx^3$	$1+(n+4)x+2x^2+3nx^3$
$(2:(1):n)$	$x+nx^3$	$1+(n+3)x+(2n+1)x^2+nx^3$
$(3:(2):n)$	$x+4nx^3$	$1+(n+5)x+4(n+1)x^2+4nx^3$

We now recall that equivalence also means "equality with respect to a certain property". In Table 7 the number of conjugated circuits [21] is computed for a homologous series of branched benzenoids. It can easily be interrred from the Table that.

$$R_3(B_j) = R_4(B_{j>=1}) = R_5(B_{j+2}) = \cdots\cdots$$

where T_j is a conjugated circuit containing (4i+2) pi-electrons.

Table 7: Conjugated circuits, R_i; for a homologous set of benzenoids whose dualits is: 

n	R_3	R_4	R_5	R_6	R_7	R_8	R_9
1	50	24	12	8	2		
2	64	50	24	12	8	2	
3	96	64	50	24	12	8	2
4	128	96	64	50	24	12	8
⋮							

2.3 FIBONACCI GRAPHS [36]

A very special type of homologous set of graphs of specific topological construction which was recently discovered [36] is called Fibonacci graphs. The simplest two such classes are being paths and cycles. More general constructions are shown in Fig. 15

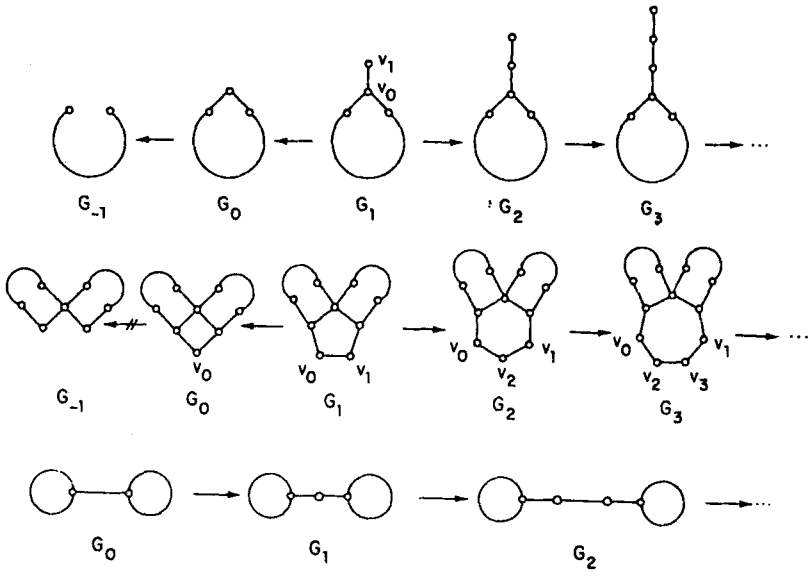


Fig. 15: External (top) and internal (bottom) construction of Fibonacci graphs.

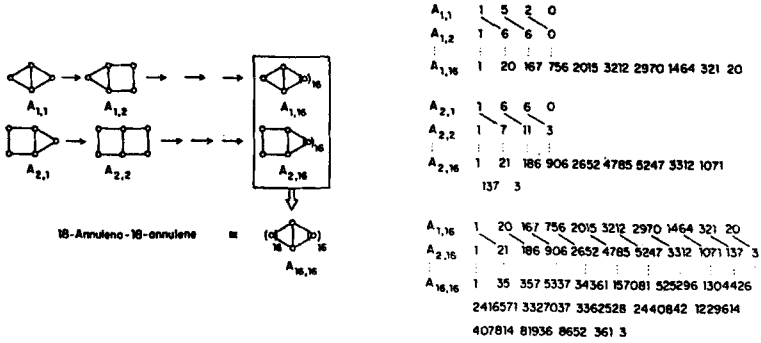


Fig. 16: Generation of the sequence of nonadjacent numbers of 18-annuleno-18-annulene by starting three fibonacci internal subdivisions.

These graphs could be of great computational importance as is illustrated in Fig. 16 where the nonadjacent numbers of 18-annuleno-18-annulene are hand-computed using the concept of Fibonacci graphs. It is worthy of mention that an attempt to obtain these nonadjacent numbers on a DEC 1099 system was unsuccessful (even after 30 minutes) [59].

PART III

APPLICATIONS

3-1. FRACTAL DIMENSION OF GRAPHS

In recent years many different graph invariants have been proposed for the characterization of chemical species which are referred to in the chemical literature as topological indices [60]. These topological parameters reflect sizes, shapes and degrees of branching of the relevant molecular graphs of the molecules of interest. In order to study the fractal nature [61] of a set of equivalent graphs such as those mentioned in PART II, one attempts to correlate a given graph invariant of a homologous series of graphs with an invariant of another sereis of graphs belonging to the same equivalence class. The situation can be formalized as follows:

Let $G_1(n) \in \underline{1}$; $G_2(n+1) \in \underline{2}$; $G_3(n+3) \in \underline{3}$, ... be a homologous series of graphs belonging to equivalence classes $\underline{1}$, $\underline{2}$, $\underline{3}$, ... Let $G'_1(n)$ $\underline{1}$, $G'_2(n+1) \in \underline{2}$, $G'_3(n+2) \in \underline{3}$, ... be another such series. The question to be posed then is the following: Is there a relation R such that:

$\chi(G_j(n+j-1))$ R $\chi(G'_j(n+j-1))$?, where $\chi(G)$ is some topological index of G. Indeed such relations do exist. As an illustration we let $\chi(g)$ be the molecular connectivity index of Randić [62], defined by

$$\chi(G) = \Sigma (d_i d_j)^{-\frac{1}{2}}$$

where d_i is the degree of vertex i and the summation is taken over all adjacent $(v_i v_j)$'s $\in G$, i.e., over all edges in G

Fig. 17 is a plot of $\chi(A)$'s vs. $\chi(B)$'s of the corresponding equivalent polyhexes for a series of benzenoid hydrocarbons possessing one kink, i.e., a hexagon with angular annelation

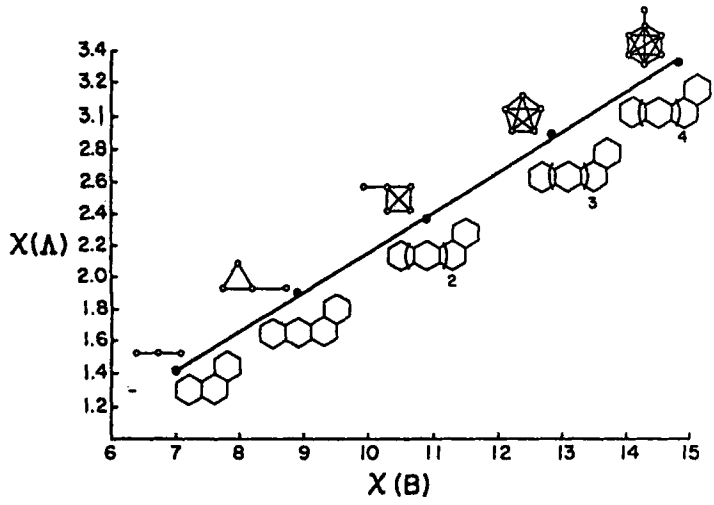


Fig. 17: A plot of molecular connectivity index, χ(B)'s, of a homologous series of polyhex graphs against the corresponding quantities, χ(A)'s, of the corresponding Clar graphs.

Fig. 18 illustrates another interesting case where the natural logarithms of the total number of acyclic Sachs graphs [63] are plotted against connectivity indices up to sixth order for a series of benzenoid hydrocarbons.

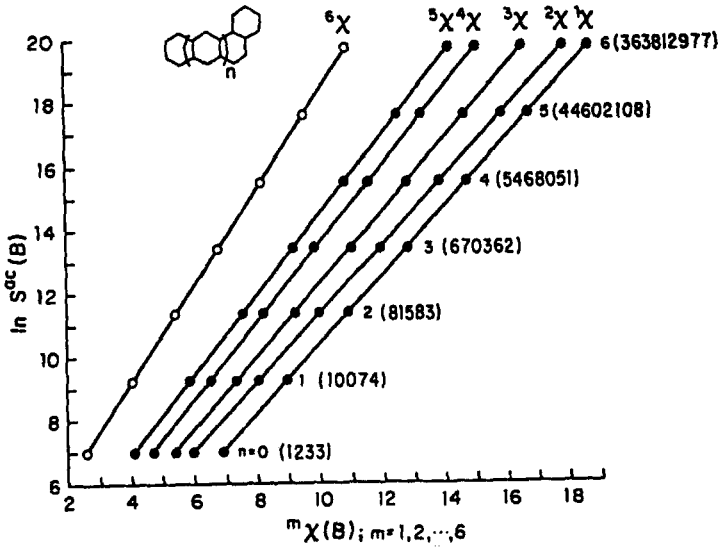


Fig. 18: Correlation between natural logarithm of the total number of acyclic Sachs graphs, ln S^{ac}(B) (numbers in parentheses) for a series of polyhex graphs with their connectivity indices, $^m\chi$ (B), up to sixth order, where

$$^m\chi\,(B) = \Sigma\,(d_i d_j \cdots\cdots d_{m+1})^{-\frac{1}{2}}$$

Fig. 18 illustrates a case where two different graph theoretical properties are plotted for the same homologous series of graphs. Fig. 19 is a more involved relation: it shows a plot of ln K(B)'s, i.e., the natural logarithm of the Kekulé counts of a homologous series of polyhex graphs versus χ (T)'s of the corresponding equivalent caterpillars for the zigzag polyacenes.

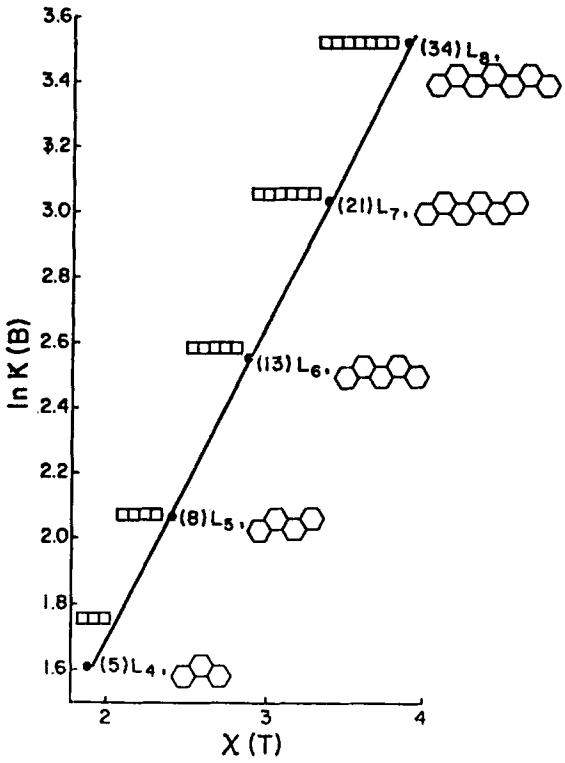


Fig. 19: A plot of logarithms of Kekulé counts, ln K(B)'s of a homologous series of benzenoids and the connectivity indices of the corresponding caterpillar trees (L_4-L_8). The equivalent king boards are drawn to the left of the correlation line. The numbers in parentheses are the individual Kekulé counts.

Since the Clar graphs equivalent to these classes are simple paths a linear relation with an identical slope results for this series when ln K(B)'s are correlated with $\chi(\Lambda)$'s. It is useful to recall here that: $K(B) = H(T;1) = \sigma(B;1) = \omega(X;1) = K(P;1) = R(P_r;1)$.

Hence Fig. 19 represents, in fact five correlations! Other many more relations are available [64]. Here, in an analogy with self-similarity the concept of fractal dimensions has been extended to graphs by

viewing members of a family as self-similar and associating with such homologous families a "fractal" dimension. Perhaps all this may sound arbitrary, yet, on the other hand, the minimal mathematical formalism is satisfied and maintained. This then justifies extension of the concept of fractals to graphs, the concept which has been already useful for other geometrical objects [65].

3.2 CORRELATION WITH PHYSICAL PROPERTIES

Sometimes a correlation exists between two different physical properties of the same compound. An example is the correlation found between boiling points and chromatographic elution times [66]. To test our equivalence relations further, one hopes to find a correlation between the physical property of a molecule (corresponding to a molecular graph) and a graph theoretical index of another graph (or molecule) which is in the same equivalence class. Several such relations have been found [64] but for the sake of illustration we consider Fig. 20: a plot of $\ln \lambda_\beta$, the natural logarithm of the β absorption band [67] of the U.V. spectra of a triphenylene homologues against connectivity indices of the equivalent pseudo-caterpillar trees.

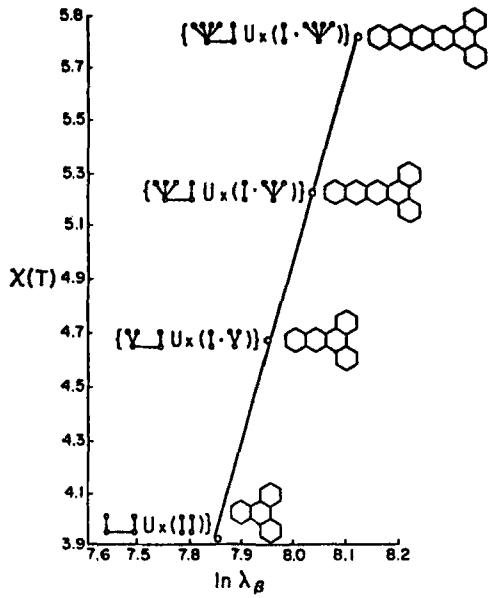


Fig. 20: A plot of electronic absorption β band, λ_β, for a series of branched benzenoids and the connectivity indices of the equivalent "pseudo" trees.

Other properties including heats of atomization magnetic properties and even reactivity towards certain cyclization reactions are also available [64].

3-3 ORDERING OF STRUCTURES; ON Randić -Wilkins Templates [25]

In a series of papers Randić and Wilkins proposed a new method to order graphs by considering their molecular path codes. Their rule of ordering states that two structures can be compared only if $p_2 \geq p'_2$ and $p_3 \leq p'_3$ where p_i is the number of self-avoiding walks of length i in a given graph, $p_i \in G_i$. This rule was applied to several alkanes, and the derived templates led to a partial ordering of several physical properties of alkanes. The grid of octanes [25] is shown in Fig. 21. The octane molecular graphs are all caterpillar trees except three (corresponding to points 9, 13 and 16 on the template of Fig. 21). In order to test the concept of equivalence classes of graphs and extend it to molecules we replace the points corresponding to caterpillar graphs on a given template by the molecules of the corresponding equivalent graphs. In Fig. 21 polyhex graphs (of benzenoid hydrocarbons) are studied in this way. The numbers corresponding to vertices of the octane grid represent mean information theoretic indices, $\bar{I}$'s [68] on the permutation integrals [12] of the benzenoid hydrocarbons. Thus let $\{\gamma_1, \gamma_2, \cdots \gamma_m\}$ be the permutation integrals involving 6-, 10-, ..., (4m + 2) pi-electrons in the benzenoid system (this sequence corresponds to $\{R_1, R_2, \cdots\cdots, R_m\}$ conjugated circuits [21] in the corresponding polyhexes), then

$$\bar{I} = (N \ln N - \sum_{i}^{m} \gamma_i \ln \gamma_i)/N$$

where N = total number of permutation integrals in the benzenoid hydrocarbon. It is clear from Fig. 21 that in all cases $\bar{I}$ increases on going from left to right and from top to bottom with only two discrepencies (points 6 and 10). Similar grids are available for higher members and include several other graph-theoretical properties.

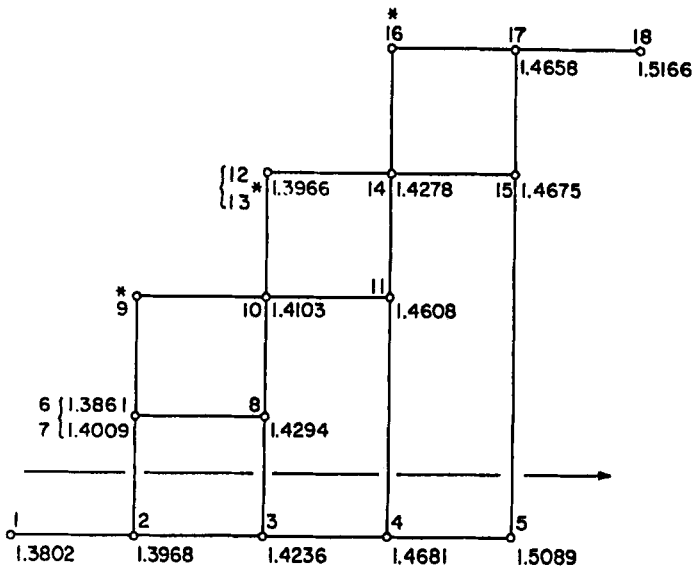


Fig. 21: A grid resulting when the Randić -Wilkins' partial ordering conditions is adopted for the isomers of octanes. The values next to the vertices are mean information-theoretic indices, $\bar{I}$, on the permutation electronic integrals of the equivalent polyhexes. Starred positions indicate cases where the molecular graph of a given octane is not a caterpillar and hence for which a polyhex is not defined. The template leads to excellent partial ordering where $\bar{I}$ values always increase from left to right.

References and Notes

1. I.L. Finar, Organic Chemistry, vol. 2, p. 417 Longmans (1968).
2. A.T. Balaban and F. Harary, Tetrahedron 24, 2505 (1968).
3. The name caterpillar was suggested by A. Hobbs; C.f., F. Harary and A.J. Schwenk, Discrete Math. 6, 359 (1973). It is a graph whose derivative is a path. See, S. El-Basil, J. Math. Chem., 1, 153 (1987).
4. I. Gutman, Z. Naturforsch a 37, 69 (1982); I. Gutman and S. El-Basil, Z. Naturforsch a, 39, 276 (1984).
5. I. Gutman, Theoret. Chim. Acta 45, 309 (1977).
6. A.T. Balaban, Tetrahedron, 25, 2949 (1969).
7. S. El-Basil, Topics in Curr. Chem., 153, 274 (1990).
8. K. Balasubramanian, Studies in physical and theoretical chemistry 23, 149 (1982).
9. R.P. Grimaldi, Discrete and Combinatorial Mathematics An applied introduction, p. 129; Addison-Wesley (1985).
10. I. Gutman and M. Randić, Chem. Phys. Lett., 47, 15 (1977).
11. F. Harary, Graphy Theory, Addison-Wesley (1969).
12. W.C. Herndon and M.L. Ellzey, Jr., J. Am. Chem. Soc., 96, 6631 (1974).
13. D.J. Klein, G.E. Hite, and T.G. Schmalz, J. Comput. Chem., 7(4), 443 (1986).
14. The name factor graph was suggested by Randić , these graphs were elaborated by one of the authors: S. El-Basil, Internat. J. Quantum Chem., 21, 771; 779; 793 (1982); Match., 13, 183: 199; 209 (1982); Chem. Phys. Lett., 89, 145 (1982); Bull. Chem. Soc. Japan, 56, 3158 (1983); Croat. Chem. Acta, 57, 47 (1984).
15. V. Krishamurthy, Combinatorics: Theory and Applications, E. Horwood, Halsted Press (1986).
16. If no three hexagons have a common atom, the system is called catacondensed. If every hexagon of a catacondensed system has at most two neighboring hexagons, it is called unbranched. More detailed notation may be found in: S.J. Cyvin and I. Gutman, Lecture Notes in Chemistry, 46 (1988).
17. Line graphs are defined in: F. Harary, Graph Theory Chapt. 8, Addision-Wesley, (1969).
18. D.I.A. Cohen, Basic Techniques of Combinatorial Theory, p. 134, John Wiley & Sons (1978).
19. A. Motoyama and H. Hosoya, J. Math. Phys. 18, 1485 (1985); S. El-Basil, J. Comput. Chem., 8, 956 (1987).
20. N. Trinajstić , Chemical Graph Theory, Vol. I, II, CRC Press, Boca Raton, Florida (1983).
21. M. Randić , Chem. Phys. Lett. 38, 68 (1976); Tetrahedron 33, 1905 (1977); J. Am. Chem. Soc., 99, 444 (1977).
22. N. Christofides, Graph Theory, An Algorithmic Approach, Academic Press, Chap. 3 (1975); S. El-Basil and M. Randić , J. Math. Chem., 1, 290, (1987); S. El-Basil, Discrete Appl. Math., 19, 145 (1988).

23. See, e.g., S. El-Basil, J. Chem. Soc., Faraday Trans, 3, 82, 299 (1986).
24. H. Hosoya, Bull. Chem. Soc. Japan 44, 2332 (1971).
25. M. Randić , G.M. Brissey, R.B. Spencer and C.L. Wilkins Computers and chemistry, 3, 5 (1979); M. Randić and C.L. Wilkins, Chem. Phys. Lett., 63, 332 (1979); J. Phys. Chem., 83, 1525 (1979); Internat. J. Quantum Chem., 18, 1005 (1980), M. Randić and N. Trinajstić , Match., 13, 271 (1982).
26. M. Randic', J. Comput. Chem., 1(4), 386 (1980).
27. S. El-Basil, Theoret. Chim. Acta, 65, 191 (1984).
28. Here the term data reduction means studying the properties of a "large" molecule in terms of those of a "smaller" one. In the past this involved mainly curve fitting procedures. The role of graph theory was recognized in the following papers: E.A. Smolenskii, Russ. J. Phys. Chem. (Engl. Transl.); 38, 700 (1964); M. Gordon and J.W. Kennedy, J. Chem. Soc., Faraday Trans. 2, 68, 848 (1972).
29. M. Randić , Chem. Phys. Lett., 55, 547 (1978); M. Randić and C.L. Wilkins, Chem. Phys. Lett., 63, 332 (1979); M. Randić , Chem. Phys. Lett., 58, 180 (1978); J. Chromatography, 161, 1 (1978); M. Randić and B.C. Gerstein, J. Magnetic Resonance, 43, 207 (1981); M. Randić , Internat. J. Quantum, Chem., 23, 1707 (1983); M. Randić , N. Trinajstic , Theoret. Chim. Acta, 73, 233 (1988).
30. E. Ruch, Theoret. Chim. Acta, 38, 167 (1975).
31. E. Ruch and A. Schonhofer, Theoret. Chim. Acta, 19, 225 (1970).
32. W. Kaplan, Advanced Calculus, p. 169, Addison-Wesley (1973).
33. S. El-Basil, Chem. Phys. Lett., 89, 145 (1982); S. El-Basil and I. Gutman, Chem. Phys. Lett., 94, 188 (1983).
34. S. El-Basil, Bull. Chem. Soc. Japan, 56, 3152 (1983).
35. I. Gutman, Croat. Chem. Acta, 56, 373 (1983).
36. S. El-Basil, J. Math. Chem., 2, 1 (1988); Theoret. Chim. Acta, 65, 191; 199 (1984).
37. N. Ohkami and H. Hozoya, Theoret. Chim. Acta, 64, 153 (1983).
38. H. Sachs, Combinatorica 4(1), 89 (1984).
39. $\sigma(B, 1)$ is the value of the sextet polynomial, $\sigma(B; x)$ for a benzenoid system B when x=1. Sextet polynomials were first introduced in: H. Hosoya and T. Yamaguchi, Tetrahedron Letters, 52, 4659 (1975).
40. $\xi(B; 1)$ is the value of the Clar polynomial $\xi(B; x)$ for x=1, c.f., S. El-Basil, Theoret. Chim. Acta, 70, 53 (1986); S. El-Basil, Discrete Appl. Math., 19, 145 (1988), S. El-Basil and M. Randić , J. Math. Chem., 1, 281 (1987).
41. S.J. Cyvin and I. Gutman, Lecture Notes in Chemistry, 46, 25-32 (1988).
42. S. El-Basil and M. Randić , J. Chem. Soc. Faraday Trans. 2, 84, 1875 (1988).

43. P. John and H. Sachs, in: Graphen in Forschung und Unterricht, eds. R. Bodendiek, H. Schumacher and G. Walther (Franzbecker-Verlag, Kiel), pp 85-101, (1985). For a "chemical version" of the "The John-Sachs Determinant" see., I. Gutman and S.J. Cyvin, Chem. Phys. Lett., 136, 137 (1987).
44. A.N. Philippou, G.E. Bergum and A.F. Horadam, eds., The Fibonacci Numbers. Applications of Fibonacci Numbers in the topological theory of, benzenoid and related graphs are reviewed in: S. El-Basil and D.J. Klein, J. Math. Chem., 3, 1 (1989).
45. G. Chartrand, Graphs as Mathematical Models, Prindle, Weber & Schmidt, Boston, Ma, Chapt. 5 (1977).
46. M.J.S. Dewar, The Molecular Orbital Theory of Organic Chemistry. McGraw-Hill (1969).
47. M. Randić , Math/Chem/Comp., 54, 101 (1987) Internat. J. Quantum Chemistry; Quantum Biology Symposium 15, 201 (1988).
48. E. Clar, The Aromatic Sextet, Wiley (1972). For a review see, I. Gutman, Bull. Soc. Chim. Beog. ad, 47, 465 (1982).
49. R.C. Read, Some recent results in chemical enumeration in: Graph "Theory and Application Eds. Y. Alavi, D.R. Lick and A.T. White, Springer Verlag, Berlin (1972).
50. C.L. Liu, Elements of Discrete Mathematics, McGraw Hill p. 59 (1977).
51. See, A.V. Aho, J.E. Hopcroft, and J.D. Ullman, The Design and Analysis of Computer Algorithms, Addison-Wesely, Reading, p. 180 (1974).
52. By "Graph Invariant" is meant here the selected sets of invariants shown in Table 3.
53. I.e., two nonisomorphic graphs which possess identical characteristic polynomials, C.f., A.T. Balaban and F. Harary, J. Chem. Docum. 11, 258 (1971).
54. Details are available on request.
55. The term "pseudo" king polyomino is used here to indicate that they are not proper polyomino graphs [37] but contain factors of the variable x. The same terminology might be extended to caterpillars, See, S. El-Basil, J. Math. Chem., 1, 161 (1987); D.H. Rouvray and S. El-Basil, J. Molecular Structure (Theochem.), 165, 13 (1988).
56. Each Kekulé valence-bond structure corresponds to a perfect match of the edges of the polyhex graph, see ref. 38.
57. J. Riordan, An Introduction to Combinatorial Analysis Wiley, New York (1958); C.L. Liu, Introduction to Combinatorial Mathematics, McGraw-Hill, New York (1968). A chemical version on rook boards may be found in C.D. Godsil and I. Gutman, Croat. Chem. Acta. 54, 53 (1981).
58. G.H. Hardy, J.E. Littlewood and G. Polya, Inequalities, Cambridge University Press, London p. 44 (1934).
59. B.A. Hess, Jr., L.J. Schaad and I. Agranant, J. Am. Chem. Soc., 100, 5268 (1978).

60. A.T. Balaban, Theoret. Chim. Acta, 53, 355 (1979); A.T. Balaban, I. Motoc, D. Bonchev, and O. Mekenyan, Top. Curr. Chem., 114, 21 (1983).
61. B.B. Mandelbrot, The Fractal Geometry of Nature, Freeman, San Francisco (1982).
62. M. Randić, J. Am. Chem. Soc., 97, 6609 (1975).
63. H. Sachs, Publ. Math. (Debrecen) 11, 119 (1964) For a "Chemical Version" see N. Trinajstić, Croat. Chem. Acta, 49, 393 (1977).
64. S. El-Basil and M. Randić, unpublished results (available on request).
65. D.H. Rouvray and R.B. Pandey, J. Chem. Phys. 85, 2286 (1986).
66. K. Kovats, Z. Anal. Chem. 181, 351 (1961).
67. Data are collected from E. Clar, The Aromatic Sextet, Wiley, New York, 101, 1 (1972).
68. D. Bonchev and N. Trinajstić, J. Chem. Phys., 67, 4517 (1977).

Index

D

E

F

G

H

I

O

P

Q

R

S

T

W

Y