PROGRESS IN THEORETICAL ORGANIC CHEMISTRY

VOLUME 1

THEORY AND PRACTICE OF
MO CALCULATIONS ON ORGANIC MOLECULES

I.G. CSIZMADIA

PROGRESS IN THEORETICAL ORGANIC CHEMISTRY

VOLUME 1

THEORY AND PRACTICE OF
MO CALCULATIONS ON ORGANIC MOLECULES

I.G. CSIZMADIA

Department of Chemistry
University of Toronto
Toronto, Ontario
Canada M5S 1A1

PREFACE

There is only one way to learn something .... by doing
it. This means that the road to knowledge in theoretical
organic chemistry is built from quantum chemical computations
on organic molecules. It is hoped that after an experimental
organic chemist has _finished_ this book he will be ready to
begin. Even if he feels he is not ready, he should start and
after some computational experience has been gained the con-
tent of this introductory book will mean considerably more.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

SECTION A

INTRODUCTION

CHAPTER I

INTRODUCTORY REMARKS

1. The Role of Theories and Models

1.   <u>The Role of Theories and Molecules</u>

        The following scheme encompasses in a very general
way all scientific activity

MODEL

Conceptual
Explanation

Theoretical
Observation

Experimental
Observation

THEORETICAL
TECHNIQUE

EXPERIMENTAL
TECHNIQUE
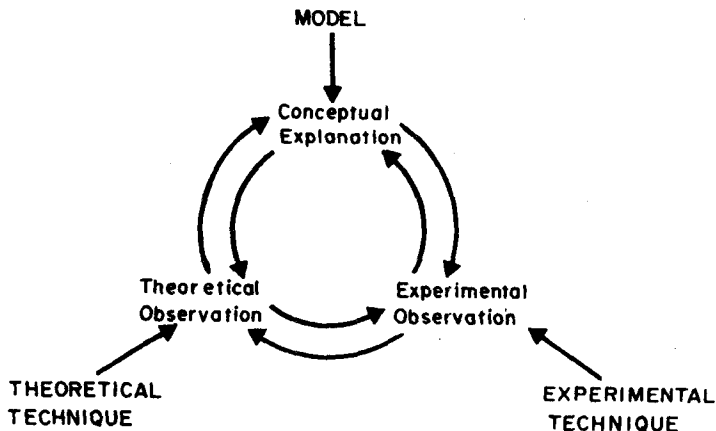
        Figure I-1.   A schematic illustration of the
    interrelationship of experiment, theory and modelling.

        It should be noted that THEORY and EXPERIMENT are
equally fundamental in any branch of science that has reached
a rigorous stage.   MODELS on the other hand are built either on
theoretical and/or experimental observations and are expected
to provide a conceptual explanation for the phenomenon
investigated.
        When its chemical implications are sought we can
identify the "THEORETICAL TECHNIQUE" with "Quantum Theory"
including all its various branches.   The "MODEL" may be iden-
tified with all the "Rules" we have in chemistry such as the
"Selection Rules", the "Woodward-Hoffmann Rules" and the like.
The application of any THEORETICAL TECHNIQUE to a particular
chemical problem is carried out through computation which in
turn produces the Theoretical (i.e. numerical) Observations in
very much the same way as the application of an EXPERIMENTAL
TECHNIQUE leads to qualitative (i.e. non-numeric) or quanti-
tative (i.e. numerical) results.   The application of a MODEL

to the same chemical problem produces qualitative (i.e. non-numerical) results which may be regarded as concepts or conceptual explanations and possessing them we usually declare that we have some understanding of the problem.

There are at least three things to be noticed. One is that the results obtained from theory are always quantitative   from models always qualitative, while experimental results may be either qualitative or quantitative. The second thing to notice is that in both theory and models we are dealing with intellectual constructs while in any chemical experiment one is working with actual chemical substances. The third point is that there exists a whole spectrum of intellectual constructs between theory and models and it need not be trivially obvious into which category a construct fits.

At this stage it is necessary to point out that the above viewpoint, in which all of our scientific activities are unified on an equal basis, by no means enjoys uniform acceptance in the chemical community. In fact, we may classify most of the confessed opinions into the following four categories.

1. Only MODELS and therefore Conceptual Explanations are of any importance and therefore of any real relevance to chemistry because THEORY is largely incomprehensible and consequently irrelevant.

2. THEORY is important in so far as it supports a useful MODEL.

3. THEORY is of primary importance and a MODEL is acceptable only in so far as it is in close agreement with the Theoretical Observations.

4. THEORY is the only scientifically acceptable and therefore relevant method that complements EXPERIMENT. Consequently all MODELS are more or less useless and thus irrelevant.

Perhaps most experimentalists' point of view could be fitted in the first three categories while most theoreticians' views could be accommodated by the last three. To the average chemist, however, the relatively moderate views of (2) and (3) are most appealing and the first and last statements appear extreme; the first one being at the ultra conservative extreme

while the last is at the radical extreme. At this stage it may
be appropriate to make an operational distinction between
theory and a model. Such a distinction may be made in terms of
their applicability. If we apply a Theory to investigate a
question that asks WHAT we may compute an answer within any
quantitative theory. For example: WHAT is the most stable
geometry of NH_3? We may obtain r(N-H) = 1.01 $\overset{o}{A}$ and <HNH = 107°.
Alternatively we may ask: WHAT is the barrier to pyramidal
inversion in NH_3? We may compute 6 Kcal/mole. However no
theory can answer a question: WHY is ammonia pyramidal or WHY
is its barrier to pyramidal inversion 6 Kcal/mole? As soon as
we try to obtain an answer to this type of question, wittingly
or unwittingly we are using a model (cf. Figure I-2.).

WHAT? ——— | Theory | ——→ An Answer
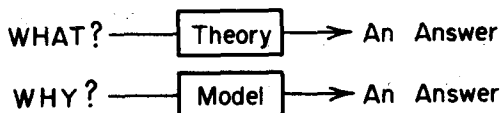
WHY? ——— | Model | ——→ An Answer

Figure I-2. The differing utility of a theory and a model.

The fundamental problem is that with the question WHY we
would like to know what is causing what. For example we may
say that NH_3 is pyramidal because such a geometry is guaranteed
by its electron distribution. Immediately then we may ask WHY
is the electron density so distributed. Then one might suggest
that the position of the nuclei is causing such an electron
distribution. Even at this stage we may ask WHY is the balance
between nuclear and electronic forces such that the pyramidal
geometry is favoured. In other words a question WHY leads to
an infinite number of questions until one arrives at the ulti-
mate or "primary cause". In contrast to this situation once an
answer is given to a question WHAT no further question of WHAT
may follow. Consequently people in category No. 4 are of the
opinion that WHAT constitutes a scientific question while WHY
constitutes a philosophical or theological question.

After summarising the rationale behind the extreme viewpoint (4), it is enough to say that quite likely you, the reader, may identify your point of view with one of the four (1-4) statements above. Irrespective of your point of view you probably wish to find out more about the THEORETICAL TECHNIQUES used for the generation of Theoretical Observations which is the subject of the subsequent chapters.

6

CHAPTER II

MATHEMATICAL INTRODUCTION

It is advantageous to treat orbitals as vectors and to deal with the mathematical problem of generating molecular orbitals (MO) from atomic orbitals (AO) as a transformation from one set of vectors to another. In this sense molecular orbital theory may be viewed as an exercise in Linear Algebra and this presentation of molecular orbital theory therefore begins with a review of vectors, vector spaces and their associated manipulations.

1. Vectors and Vector Spaces

A vector space is a set of mathematical objects ϕ, χ, η... called vectors. Although we may think of a vector ϕ as a geometrical object characterized by a length and a direction, however, it is best to abandon this fixed idea because we shall consider other than geometrical objects, such as mathematical functions to be vectors.

$$\phi = \frac{1}{\sqrt{\pi}} e^{-\sqrt{x^2 + y^2 + z^2}}$$
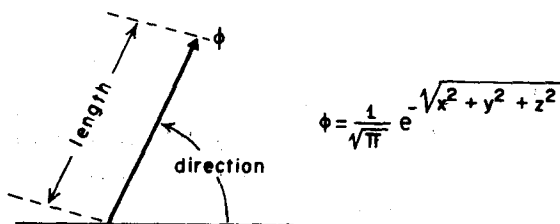
Figure II.1   The vector as geometrical object and as mathematical function.

Consequently, if we wish to consider abstract vector spaces, which is the most convenient way to discuss quantum chemistry then it is best to consider vectors as abstract mathematical objects. However, from time to time geometric illustrations will be given to facilitate understanding.

For any pair of vectors that belong to the vector space there is a unique sum The addition of vectors is both commutative and associative

$$\phi + \chi = \chi + \phi \qquad \text{(commutative)} \qquad \{II-1\}$$

$$\phi + (\chi + \eta) = (\phi + \chi) + \eta \qquad \text{(associative)} \qquad \{II\text{-}2\}$$

There is one unique vector in every vector space called origin and zero vector: O and for every vector ϕ there exists an inverse vector $-\phi$. These special vectors are subject to the following addition laws

$$\phi + O = \phi \qquad\qquad \{II\text{-}3\}$$
$$\phi + (-\phi) = O \qquad\qquad \{II\text{-}4\}$$

For each scalar a and each vector ϕ there exists a multiple vector which is the product of a and ϕ. The multiplication of scalars is both distributive and associative

$$a(\phi + \chi) = a\phi + a\chi \qquad\qquad\qquad \{II\text{-}5\}$$
$$(a + b)\phi = a\phi + b\phi \quad \text{(distributive)} \qquad \{II\text{-}6\}$$
$$(ab)\phi = a(b\phi) \quad \text{(associative)} \qquad \{II\text{-}7\}$$

The minimum number of vectors required to define the vector space is termed the dimension of the vector space. It is easy to provide geometrical illustrations of one, two and three dimensional vector spaces.
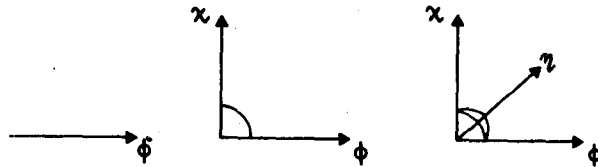

Figure II.2 Geometrical illustration of 1, 2 and 3 dimensional vector spaces.

To define a vector space it is necessary to have a set of linearly independent vectors. If the equation

$$k_1\,\eta_1 + k_2\,\eta_2 + \ldots\ldots + k_n\,\eta_n = 0 \qquad \{II\text{-}8\}$$

is valid only for <u>all</u> k_i = 0 the set of vectors $\{\eta_i\}$ <u>is</u>
<u>linearly independent</u> (if any $k_i \neq 0$ the vectors are linearly
dependent).

A linearly independent set of vectors which spans a
vector space represents the <u>basis of the vector space</u>. Any
other vector in this vector space is linearly dependent on the
basis set (i.e. not all k_i = 0) and may therefore be expressed
as a <u>linear combination of the basis vectors</u>

$$\phi = C_1\eta_1 + C_2\eta_2 + \ldots + C_n\eta_n = \sum_{i=1}^{n} C_i\eta_i \qquad \{II-9\}$$

where the set of numbers c_1, c_2 .... c_n are the <u>components</u> of
the vector over the <u>basis set</u> $\{\eta\}$.

A geometrical illustration of this principle for a two
dimensional vector space is given in the following figure.

$$\phi = (\cos\alpha)\eta_1 + (\sin\alpha)\eta_2$$
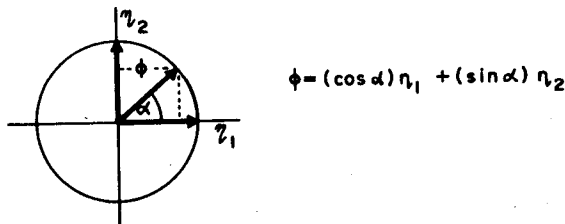
Figure II.3   The concept of linear dependence in
a two dimensional vector space.

Since it is impossible for both components ($\sin\alpha$ and $\cos\alpha$) to
be zero at any value of α, ϕ is linearly dependent on the set
of $\{\eta\}$ which is <u>the basis set of the vector space</u>.

The basis vectors may be arranged in a row (row vector)
and abbreviated as $<\eta|$

$$<\eta| = (\eta_1, \eta_2 \ldots , \eta_n) \qquad \{II-10\}$$

or in a column (column vector) and denoted as $|\eta>$

$$|\eta> = \begin{pmatrix} \eta_1 \\ \eta_2 \\ . \\ . \\ . \\ . \\ \eta_n \end{pmatrix} \qquad \{II-11\}$$

The transpose (denoted by a prime) of a column vector is a row vector:

$$|\eta>' \equiv \begin{pmatrix} \eta_1 \\ \eta_2 \\ . \\ . \\ . \\ . \\ \eta_n \end{pmatrix}' = (\eta_1 \eta_2 \ldots \ldots \eta_n) \equiv <\eta| \qquad \{II-12\}$$

and <u>vice versa</u>.

If the basis vectors are not real but complex (such as complex functions) the corresponding relationship holds for the adjoint (denoted by a dagger) which is the transpose complex conjugate (the complex conjugate is denoted by an asterisk)

$$|\psi>^{\dagger} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ . \\ . \\ . \\ . \\ \psi_n \end{pmatrix}^{\dagger} = (\psi_1^{\star} \psi_2^{\star} \ldots \ldots \psi_n^{\star}) \equiv <\psi| \qquad \{III-13\}$$

With this notation at hand it is possible to develop a vector notation for the linear combination of the basis vectors as

defined in equation {II-9}

$$\phi = c_1\eta_1 + c_2\eta_2 + \cdots + c_n\eta_n = (c_1 c_2 \cdots c_n) \begin{pmatrix} \eta_1 \\ \eta_2 \\ \cdot \\ \cdot \\ \cdot \\ \eta_n \end{pmatrix} \quad \{II-14\}$$

or in its equivalent form

$$\phi = \eta_1 c_1 + \eta_2 c_2 + \cdots + \eta_n c_n = (\eta_1 \eta_2 \cdots \eta_n) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad \{II-15\}$$

It probably should be pointed out at this stage that this is exactly the method (as indicated in equations {II-9}, {II-11} and {II-15} used to generate molecular orbitals from atomic orbitals where the basis set {η} stands for the set of AO used and ϕ represents one particular molecular orbital. The **expansion coefficients** (i.e. the coefficients of the linear combination) are labelled as $c_1 c_2 \cdots c_n$ and these are in fact the **components** of the particular MO: ϕ over the chosen AO basis {η}.

2. Inner Product and Orthogonality

Molecular orbital theory involves special types of vector spaces where the basis set (the atomic orbitals) consists of scalar-valued continuous functions of space (i.e. x, y, z) defined in an interval {a,b} for each one of the independent variables (x, y, z). This example may be referred to as a vector space of continuous functions on {a,b}. It may be noted that using a Cartesian coordinate system the interval is {$-\infty$, $+\infty$} for all three independent variables.

For such a vector space the inner product of any two basis vectors is defined as the following definite integral

$$S_{ij} = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \eta_i(x,y,z)\eta_j(x,y,z) \; dx \; dy \; dz \qquad \{\text{II-16}\}$$

If the set of orbitals $\{\eta\}$ used are complex valued functions the above definition should be generalized as follows

$$S_{ij} = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \eta_i{}^*(x,y,z)\eta_j(x,y,z) \; dx \; dy \; dz \qquad \{\text{II-17}\}$$

This integral (which is frequently called the overlap integral) is abbreviated in the following fashion in Dirac's notation.

$$S_{ij} \equiv \langle \eta_i | \eta_j \rangle \qquad \{\text{II-18}\}$$

where the $\langle \eta_i |$ includes the complex conjugate as specified in equation $\{\text{II-17}\}$.

When S_{ij} is zero then the two vectors η_i and η_j are said to be orthogonal. Although orthogonality is a more general concept than perpendicularity, however it may be help-ful to review a geometrical example.

As indicated by the geometrical relationship shown in Figure II-4 the inner product:

$$S \equiv \langle \phi | \chi \rangle = |\phi| \cdot |\chi| \cdot \cos\theta \qquad \{\text{II-19}\}$$

will vanish when the two vectors are perpendicular to each other (i.e. when the angle θ is 90° or 270°).

In quantum chemistry orthogonalization of the basis i.e. the basis vectors (e.g. the AO) facilitates the calcul-ation and is therefore of some importance. The most frequently used methods are the symmetric (Löwdin) orthogonalization and the Schmidt orthogonalization. The geometrical illustration of these two methods of orthogonalization is shown in Figure II-5.

In actual calculations one usually works with more than two basis vectors (i.e. AO) and the whole set needs to be orthogonalized so that any pair of vectors in the set will be