Title
COMPUTATIONAL METHODS FOR MOLEUCLAR STRUCTURE DETERMINATION: THEORY AND TECHNIQUE

Author
Lester, W.A.

Publication Date
1980-03-14

Peer reviewed

CONTENTS

FOREWORD

The National Resource for Computation in Chemistry (NRCC) was established as a Division of Lawrence Berkeley Laboratory (LBL) in October 1977. The functions of the NRCC may be broadly categorized as follows: (1) to make information on existing and developing computational methodologies available to all segments of the chemistry community, (2) to make state-of-the-art computational facilities (both hardware and software) accessible to the chemistry community, and (3) to foster research and development of new computational methods for application to chemical problems.

Workshops are one facet of the NRCC's program for both obtaining and making available information on new developments in computationally oriented subdisciplines of chemistry. The goal of this workshop was to provide an introduction to the use of state-of-the-art computer codes for the semi-empirical and *ab initio* computation of the electronic structure and geometry of small and large molecules.

The workshop consisted of lectures on the theoretical foundations of the codes, followed by laboratory sessions which utilized these codes. The lectures, many of which were presented by the original developers of the computational methods, provided the participants with a knowledge of the strengths and weaknesses of the various theoretical methods. The laboratories, which were conducted by NRCC and QCPE staff, provided a unique "hands-on" experience for the participants. Through the use of remote interactive terminals and a remote job entry station, they were able to utilize all of the methods presented in the lectures in an examination of chemically interesting systems.

Forty-five participants from the academic, industrial, and governmental sectors attended this workshop.

The material contained in these proceedings consists solely of the partially edited lecture notes provided by the guest speakers. They are reproduced here to convey the essence of the subject matter covered in the lectures to those not in attendance.

The NRCC is indebted to QCPE for helping to organize this workshop, to the Indiana University Chemistry Department for making their facilities available, to the Indianapolis office of the Control Data Corporation for providing a remote job entry station, and to the Computer Science and Electronics departments at Lawrence Berkeley Laboratory for providing help and technical assistance.

We also thank Drs. Michel Dupuis and John J. Wendoloski of the NRCC for their efforts in organizing this volume.

The National Resource for Computation in Chemistry is supported in part by a grant from the National Science Foundation (Grant No. CHE-7721305) and the Basic Energy Sciences Division of the U.S. Department of Energy (Contract No. W-7405-ENG-48).

The Quantum Chemistry Program Exchange is a self-supporting organization and is part of the Department of Chemistry of Indiana University, Bloomington, Indiana.

William A. Lester, Jr.,
Director NRCC

LIST OF INVITED SPEAKERS

Professor Norman L. Allinger
Department of Chemistry
University of Georgia
Athens, Georgia

Professor Ernest R. Davidson
Department of Chemistry
University of Washington
Seattle, Washington

Dr. Bowen Liu
IBM Research Laboratory
San Jose, California

Professor William A. Goddard, III
Department of Chemistry
California Institute of Technology
Pasadena, California

Professor John Pople
Department of Chemistry
Carnegie-Mellon University
Pittsburgh, Pennsylvania

Professor Michael C. Zerner
Guelph-Waterloo Centre for
Graduate Work in Chemistry
University of Guelph
Ontario, Canada

WORKSHOP PARTICIPANTS

Suheil F. Abdulnur
National Foundation for
Cancer Research
Chemistry Department
American University
Washington, D.C. 20016
(301)686-0039

Stephen M. Adams
New York State Department
of Health
Empire State Plaza
Albany, NY 12201
(518)473-1336

Beatrice Botch
Chemistry Division
Argonne National Laboratory
Argonne, IL 60439
(312)972-4803

Donald Boyd
Eli Lilly Company
The Lilly Research Laboratories
740 S. Alabama
Indianapolis, IN 46206
(812)261-4133

James F. Caputo
College of Pharmacy
University of Iowa
Iowa City, IA 52242
(319)353-3041

B. Vernon Cheney
The Upjohn Company
Kalamazoo, MI 49001
(616)385-7805

Wai-Yim Ching
Department of Physics
University of Missouri
Kansas City, MO 64110
(816)276-1604

Ashok L. Cholli
Macromolecular Science Department
Case Western Reserve University
Cleveland, OH 44106

James R. Damewood
Department of Chemistry
Princeton University
Washington Road
Princeton, NJ 08544
(609)452-5093

Charles D. Duncan
Department of Chemistry
University of Alabama, Birmingham
Birmingham, AL 35294
(205)934-4747

Robert A. Eades
Department of Chemistry
139 Smith Hall
University of Minnesota
Minneapolis, MN 55455
(612)373-7923

W. Daniel Edwards
Department of Chemistry
University of Guelph
Guelph, Ontario N1G-2W1 Canada
(519)824-4120

Paul F. Endres
Chemistry Department
Bowling Green State University
Bowling Green, OH 43403
(419)372-0315

Gernot Frenking
Technische Universität Berlin
Institut für Organische Chemie
der TU Berlin
Strasse des 17, Juni 135
D-1000 Berlin 12, West Germany
(030)314-3621

Donald M. Friedrich
Department of Chemistry
Hope College
Holland, MI 49423
(616)392-5111, x.3222

Neil M. Goldstein
Department of Chemistry
University of Massachusetts
Amherst, MA 01003
(413)545-2534

Barry Gordon
Department of Chemistry
New York University
New York, NY 10003

mailing address:
153 Beach 118th St.
Rockaway Park, NY 11694
(212)GR4-7567

Kenneth W. Hedberg
Department of Chemistry
Oregon State University
Corvallis, OR 97331
(503)754-2371

H. Fred Henneike
Chemistry Department
Georgia State University
University Plaza
Atlanta, GA 30303
(404)658-3120

Richard J. Johnson
Department of Biochemistry
Duquesne University
Pittsburgh, PA 15243
(412)434-6340

Herbert W. Jones
Physics Department
Florida A&M University
Tallahassee, FL 32307
(904)599-3826

Richard T. Keys
Chemistry Department
California State University, L.A.
5151 State University Drive
Los Angeles, CA 90032
(213)244-3232

Zvi C. Kornblum
School of Engineering & Chemistry
The Cooper Union
Cooper Square
New York, NY 10003
(212)254-6300, x.273

Sunil K. Kunnathedathil
Department of Chemistry
Michigan State University
East Lansing, MI 48824
(517)353-1724

Munime Lundeen
P.O. Box 26123
Honolulu, HI 96825
(808)237-8500

Horace F. Martin
Rhode Island Hospital
395 Eddy Street
Providence, RI 02902
(401)277-5019

Charles W. McFarland
Rohco, Inc.
3203 West 71st Street
Cleveland, OH 44102
(216)651-7300

John M. McKelvey
Eastman Kodak Co.
B82 Research Laboratories
Rochester, NY 14650
(716)458-1000, x.73335

Douglas S. McNair
Department of Pathology & Medicine
Baylor College of Medicine
Texas Medical Center
1220 Moursund
Houston, TX 77030
(713)790-2419

Conchita Zuazaga de Ortiz
Laboratory of Neurobiology
UPR Medical Sciences Campus
Boulevard del Valle 201
San Juan, PR 00901
(809)723-4145

Alejandro Pisanty
Facultad de Quimica
Universidad Nacional de Mexico
mailing address:
Miami 78
Mexico 18 D.F., Mexico
5 48-82-10

Patricia L.M. Plummer
Physics & Cloud Physics
University of Missouri, Rolla
Rolla, MO 65401
(314)341-4340

Ronald D. Poshusta
Chemistry Department
Washington State University
Pullman, WA 99164
(509)335-3362

P. L. Prasad
Department of Chemistry
Syracuse University
Syracuse, NY 13210
(315)423-3697

Nelson G. Rondon
Department of Chemistry
Louisiana State University
Baton Rouge, LA 70803
(504)388-2985

Michael J. Rothman
Chemistry Department
University of Michigan
Ann Arbor, MI 48109
(313)764-8229

Fernando Ruette
Department of Chemistry
University of Arkansas
Carlson Terrace E-201
Fayetteville, AR 72701
(501)575-3103

Steve Scheiner
Chemistry Department
Southern Illinois University
Carbondale, IL 62901
(618)453-5721

William D. Stanbro
Applied Physics Laboratory
Johns Hopkins University
Johns Hopkins Road
Laurel, MD 20810
(301)953-7100, x.2480

Trina Valencich
Chemistry Department
California State University, L.A.
5151 State University Drive
Los Angeles, CA 90291
(213)244-3250

Ming-Yu Rachel Wang
Chemistry Department
Whitworth College
Hawthorne Avenue
Spokane, WA 99251
(509)466-1000, x.510

Philip Warner
Department of Chemistry
Iowa State University
Ames, IA 50011
(515)294-2788

Herschel J. R. Weintraub
Department of Medicinal Chemistry
School of Pharmacy
Purdue University
West Lafayette, IN 47907
(317)749-2509

Carter T. White
Naval Research Laboratory
Code 6171
Washington, D.C. 20375
(202)767-3550

Ian H. Williams
Department of Chemistry
University of Kansas
Lawrence, KS 60045
(913)864-4080

Tonny P. C. Wong
Chemistry Department
Dalhousie University
Halifax, Nova Scotia B3H 4J3
Canada

INTRODUCTION TO COMPUTATIONAL QUANTUM CHEMISTRY

Lecture 1

by

Ernest R. Davidson

Department of Chemistry
University of Washington
Seattle, Washington

The Schrödinger equation (1926) for the stationary states of a molecule

$$H\psi = E\psi$$

is intractable! Nevertheless, by 1931 a well defined set of approximations had been outlined capable of giving qualitative <u>or</u> quantitative information about the nature of the chemical bond. These approximations are detailed below.

<u>BORN OPPENHEIMER</u>

$$H_{el}\ \psi_{e\ell}(\underline{r};\underline{R}) = U(R)\ \psi_{e\ell}(\underline{r};\underline{R})$$

$$[K.E._N + U(\underline{R})]\ \psi_N(\underline{R}) = E\ \psi_N(\underline{R})$$

<u>LINEAR VARIATION METHOD</u> (Hyleraas, James-Coolidge, etc.)

$$\psi_{e\ell} = \sum C_I\ \Phi_I$$

$$\frac{\langle \psi_{e\ell}|H|\psi_{e\ell}\rangle}{\langle \psi_{e\ell}|\psi_{\ell\ell}\rangle} = \text{MIN W.R.T. } C_I$$

Φ_I arbitrary functions

INDEPENDENT PARTICLE MODEL

$$H \cong \sum_{i=1}^{N} G(i)$$

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1)\alpha(1) & \phi_1(2)\alpha(2) & \cdots & \phi_1(N)\alpha(N) \\ \phi_1(1)\beta(1) & \vdots & & \\ \phi_2(1)\alpha(1) & & & \\ \phi_2(1)\beta(1) & & & \\ \vdots & & & \end{vmatrix}$$

$=$ Slater determinant $\left(\begin{array}{c}\text{Pauli exclusion}\\ \text{principle}\end{array}\right)$

ψ = sum of products, each product has same E

$$G\phi_k = \varepsilon_k \, \phi_k$$

$$E = \sum \varepsilon_k \, n_k$$

Useful for free electron gas model of solid.

HARTREE-FOCK (ATOMS)

$$G = G(\rho)$$

Self-consistent field defined SCF atomic orbitals.

CONFIGURATION-INTERACTION:

Literally perturbative interactions between atomic configurations.

$$H^{(1)} = H - \Sigma G(i) = \text{perturbation}$$

LCAO-MO (TIGHT BINDING) APPROACH TO MOLECULES:

Linear variation method applied to

$$G\phi_i = \varepsilon_i \phi_i$$

$$\phi_i = \sum x_{ji} f_j$$

where x is coefficient and f = atomic orbital,

or STD ~ AO

$$\delta \frac{\langle \phi_i | G | \phi_i \rangle}{\langle \phi_i | \phi_i \rangle} = 0 \longrightarrow \begin{array}{l} S_{ij} = \int f_i f_j d\tau \\ G_{ij} = \int f_i G f_j d\tau \end{array}$$

$$\sum_q (G_{pq} - \varepsilon_i S_{pq}) x_{qi} = 0 \qquad \text{all } p$$

i.e., $\underset{\sim}{G}\,\underset{\sim}{X}_i = \varepsilon_i \underset{\sim}{S}\,\underset{\sim}{X}_i$

This approach is inadequate whenever more than one configuration is close in energy,

i.e., H_2

$$\phi_1 \cong (1s_A + 1s_B)\Big/\sqrt{2(1+S)}$$

$$\psi = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_1\alpha(1) & \phi_1\alpha(2) \\ \phi_1\beta(1) & \phi_1\beta(2) \end{vmatrix} = \frac{1}{\sqrt{2}} \phi_1\phi_1(\alpha\beta - \beta\alpha)$$

Large R

$$\psi_{MO} \longrightarrow \left\{\frac{1s_A(1)\ 1s_B(2) + 1s_B(1)\ 1s_A(2)}{\sqrt{2}}\right.$$

$$\left. + \frac{1s_A(1)\,1s_A(2) + 1s_B(1)1s_B(2)}{\sqrt{2}}\right\}$$

$$\times \left\{\frac{\alpha(1)\beta(2) - \beta(1)\alpha(2)}{\sqrt{2}}\right\}$$

while

$$\psi_{exact} \longrightarrow \left(\frac{1s_A 1s_B + 1s_B 1s_A}{\sqrt{2}}\right)\left(\frac{\alpha\beta - \beta\alpha}{\sqrt{2}}\right)$$

NOTE: Independent Particle model (MO theory) is intrinsically wrong at large bond lengths, i.e., at small S.

VALENCE BOND MODEL

$$\psi \cong \mathcal{A}\{\psi_A(1 \ldots N_A) \quad {}_B(N_A+1 \ldots N_A+N_B) \ldots\}$$

approaches correct asymptote $(R \to \infty)$. Ψ_A atomic wavefunction much more accurate than ψ_{MO} but very difficult to compute.

Problem:

Spin Couplings;

$$CH_4 \quad [C(sp^3)\,{}^4S] \cdot 4H(^2S)$$

8 unpaired electrons $\to$ 14 ways to make singlet.

Problem:

Ionic and other atomic configurations are very important;

$$[C(s^2p^3)\,{}^3P] \cdot 4H(^2S)$$

$$[C^-(s^2p^3)] \cdot (4H)^+$$

Problem:

Non-orthogonal CI with many electrons is intractable on computer;

(cost ~ N^4/matrix element)

<u>But</u>: semi-empirical VB "resonance" calculations were good for predicting "resonance energy," and may be better than was realized for certain excitation energies.

EQUIVALENCE OF MO/CI AND VB/CI

$$\phi_I = \mathcal{A}\{f_{i_1} \cdots f_{i_N}\}$$

$$\Psi = \sum C_I \Phi_I$$

<u>or</u>

$$\Phi_I' = \mathcal{A}\{\phi_{i_1} \cdots \phi_{i_N}\}$$

$$\Psi = \sum C_I' \Phi_I'$$

Every Φ_I' is linear combination of ϕ_I, so ψ's are the same. MO/CI is easier to carry out. Both are hard to interpret.

HÜCKEL THEORY

Borrowed from solid state.

Approximate matrix elements.

π ELECTRON THEORY

By symmetry if there is a mirror plane, there are $A'(\sigma)$ and $A''(\pi)$ orbitals

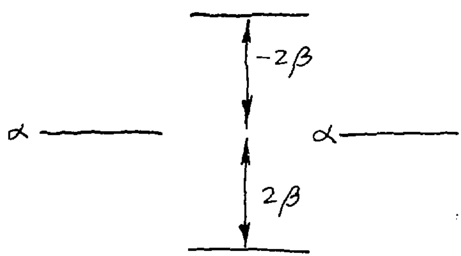
$$\pi_{MO_j} = \sum x_{ji}(f_{\pi AO})$$

$$\langle f_i | f_j \rangle = S_{ij} \cong (0.25) \cong 0$$

$$\langle f_i | G | f_j \rangle \cong \begin{cases} \alpha & i=j \\ \beta & i,j \text{ adjacent} \\ 0 & i,j \text{ not adjacent} \end{cases}$$

$$C = C \quad \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \varepsilon \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\phi_1 = \frac{1}{\sqrt{2}} (f_A + f_B)$$



$\beta < 0$

ε correlates well with

Excitation E

IP

EA

Reduction potential

$\sum \varepsilon_i n_i \leftrightarrow$ resonance energy

HETEROATOM HÜCKEL THEORY

Forced assignment of $\alpha_C - \alpha_N$, etc.

Compared to atomic SCF

$$-\varepsilon \leftrightarrow \text{IP}$$

so,

$$\alpha_C - \alpha_N \leftrightarrow (-I_C) - (-I_N) \quad .$$

Variation of bond lengths and twisting (non-planarities)

$\leftrightarrow$ variation of β

$$\beta \cong \beta^o(s/s^o)$$

i.e., $G_{ij} \sim S_{ij}$

ω method (charge self-consistency)

$$\alpha \sim \alpha^o - \omega q$$

Overlap inclusion $\underset{\sim}{G}\,\underset{\sim}{x} = \underset{\sim}{S}\,\underset{\sim}{x}\,\varepsilon$ does not particularly improve result.

EXTENDED HÜCKEL THEORY

$$\phi_i = \sum x_{ji} f_j$$

LCAO-MO $\quad f_j$ literally atomic orbital

$$G = \text{effective Hamiltonian}$$

$$S_{ij} = \langle f_i | f_j \rangle \cong \langle STO_i | STO_j \rangle$$

$\{f_i\}$ all valence orbitals

$$G_{ij} = k_{ij} S_{ij} \qquad \text{(Longuet-Higgins)}$$

$$= k_{ij} S_{ij} (G_{ii} + G_{jj})/2 \qquad \text{(Wolfsberg Helmholz)}$$

$$= k_{ij} S_{ij} \sqrt{G_{ii} G_{jj}} \qquad \text{(Ballhausen \& Gray)}$$

etc.

$$-G_{ij} = \text{valence orbital ionization potential}$$

Rotational Invariance

k_{ij} same for all orbitals with same $n\ell$ on same atom

and

H_{ij} is linearly related to S_{ij}

$$\underset{=}{G}\, \underset{-}{x} = \underset{=}{S}\, \underset{-}{x}\, \underset{=}{\varepsilon}$$

i.e.,

$$\underset{=}{G}\, \underset{-}{x}_i = \underset{=}{S}\, \underset{-}{x}_i \varepsilon_i$$

$$x_i = \begin{pmatrix} x_{1i} \\ x_{2i} \end{pmatrix} \qquad \phi_i = \sum_j x_{ji} f_j$$

$\vdots$

Non-orthogonal Eigenvalue Problem

$$H_2 \; \{1s_a \; 1s_b\} = (f)$$

$$\underline{S} = \begin{pmatrix} 1 & S \\ S & 1 \end{pmatrix}$$

$$\underline{G} = \begin{pmatrix} -I & -kSI \\ -kSI & -I \end{pmatrix}$$

$$\phi_1 = \frac{1s_A + 1s_B}{\sqrt{2(1+S)}}$$

$$\phi_2 = \frac{1s_A - 1s_B}{\sqrt{2(1-S)}}$$

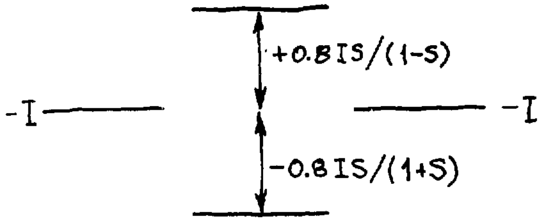
$$\varepsilon_1 = \frac{-I \; -kSI}{1 + S} = -I \left(\frac{1 + kS}{1 + S}\right)$$

$$\varepsilon_2 = -I \left(\frac{1 - kS}{1 - S}\right)$$

$$\varepsilon_1 = -I - \frac{IS(k-1)}{1 + S}$$

$$\varepsilon_2 = -I + \frac{IS(k-1)}{1 - S}$$

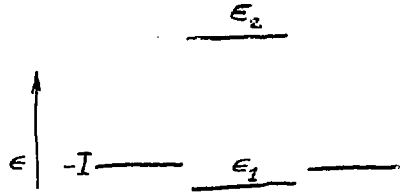
$k \sim 1.8$

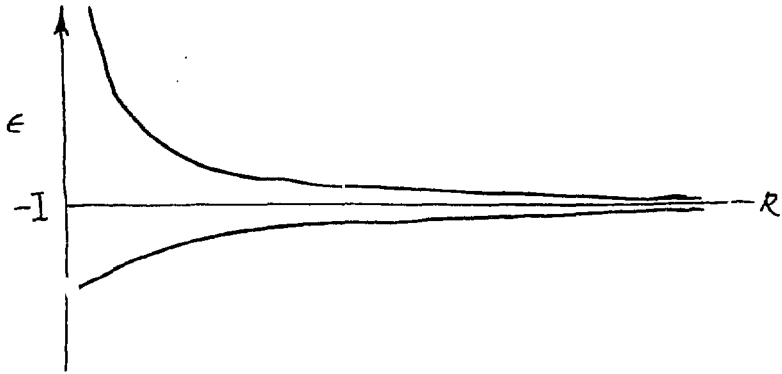


at R_e $S \sim 0.75$

$\varepsilon_2 \sim -I + 2.4I$

$\varepsilon_1 \sim -I - 0.3I$





<u>Significance of ε</u>

$$-\varepsilon(R_e) \cong I(R_e) \qquad ?$$

$$E_{\phi_1 \to \phi_2} \cong \varepsilon_2 - \varepsilon_1 \qquad ?$$

$$U(R) \cong 2\varepsilon_1 \qquad ?$$

All true for $R \geqslant R_e$. But $2\varepsilon_1$ has no minimum except for $R=0$, so it cannot be used to give geometry.

Hoffman has used $U \cong \sum \varepsilon_i n_i$ to approximate bond angles. Others have used this for bond lengths of more complicated systems with rather poor results.

Notice that $\quad 2\varepsilon \xrightarrow[R \to \infty]{} E(\infty)$

while $\quad \langle \psi | H | \psi \rangle \not\xrightarrow[R \to \infty]{} E(\infty)$

because ψ goes to wrong limit.

$$\psi = \frac{1}{\sqrt{2}} \{\phi_1 \bar{\phi}_1\} \longrightarrow (1s_A 1s_B + 1s_B 1s_A)/2$$

$$+ (1s_A 1s_A + 1s_B 1s_B)/2$$

50% covalent (correct limit)

\+ 50% ionic

<u>Evaluation of VOIE</u> (Virtual Orbital Ionization Energy):

<u>Case I</u>: Non-Iterative Neutral Molecule

$$(\mathrm{VOIE})_i \cong \text{IP of atom in promoted state}$$

i.e., for sp^3C

$$(\mathrm{VOIE})_s \cong \bar{E}(p^3C^+) - \bar{E}(sp^3C)$$

where $\bar{E}$ is configuration average.

<u>Case II</u>: Iterative EHT

Mulliken population analysis

$$N = 2\sum_{i=1}^{N/2} \int |\phi_i|^2 \, d\tau$$

$$\rho = 2\sum_{i=1}^{N/2} |\phi_i|^2 \, d\tau = \text{electron density}$$

$$\rho = \sum P_{ab} f_a f_b^*$$

$$P_{ab} = \sum_{i=1}^{N/2} x_{ai} x_{bi}$$

$$N = \int \rho \, d\tau = \sum_{a,b} P_{ab} S_{ba}$$

Define

$$q_a = \sum_b P_{ab} S_{ba}$$

$$\sum q_a = N$$

$$q_a = P_{aa} + \underbrace{\sum_{b \neq a} P_{ab} S_{ba}}_{\text{overlap population}}$$

$$Q_A = Z_A - \sum_{A \text{ on } a} q_A = \text{net charge}$$

Re-evaluate VOIE for this Q_A and this configuration by interpolation. Iterate to self-consistency

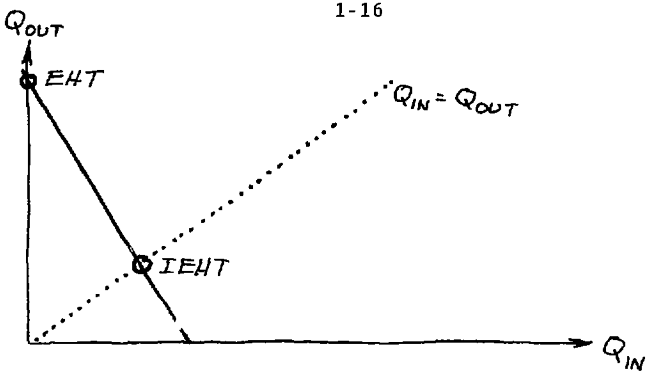
$$Q \longrightarrow \text{VOIE} \longrightarrow \text{new } Q \ .$$

Results:

IEHT gives very low net Q .

EHT gives very high net Q .

$$\begin{cases} \text{VOIE is very sensitive to Q .} \\ \text{Q is very sensitive to VOIE .} \end{cases}$$



Compared to *ab initio*, IEHT $|\eta|$ is too small

PPP (CI) with empirical integrals → good energies for vertical excitation (πe^- theory)

CNDO etc. imitate *ab initio* SCF → bad energy for right reason

MNDO
refined EHT
good energy from MO ψ without CI

or

cheap *ab initio*

LCAO-MO-SCF-CI

INTRODUCTION TO SCF THEORY

Lecture 2/3

by

Ernest R. Davidson

Department of Chemistry
University of Washington
Seattle, Washington

HARTREE-FOCK METHOD

For single determinant wavefunctions.

Closed shell Hartree Fock,

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1\alpha(1) & \phi_1\alpha(2) & \phi_1\alpha(3) \\ \phi_1\beta(1) & \phi_1\beta(2) & \vdots \\ \phi_2\alpha(1) & \vdots & \vdots \\ \phi_2\beta(1) & \vdots & \\ \vdots & & \end{vmatrix}$$

$$= \frac{1}{\sqrt{N!}} \mathrm{Det} \left\{ \phi_1\bar{\phi}_1 \; \cdots \; \phi_{N/2}\,\phi_{N/2} \right\}$$

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \qquad \text{orthonormal MO's}$$

$$E = \sum_{i=1}^{N/2} 2h_{ii} + \sum_{i,j} 2J_{ij} - K_{ij}$$

where

$$h_{ii} = \int \phi_i \; h \; \phi_i \; d\tau$$

$$h = \mathrm{K.E.} + V_{N-e}$$

$$J_{ij} = \int |\phi_i(1)|^2 \frac{1}{r_{12}} |\phi_j(2)|^2 \, d\tau_1 d\tau_2$$

$$K_{ij} = \int \phi_i(1)\phi_j(1) \frac{1}{r_{12}} \phi_i(2)\phi_j(2) \, d\tau_1 d\tau_2$$

NOTE $\quad J_{ii} = K_{ii}$

$$\left\{ \begin{array}{c} \frac{\delta}{\delta\langle\phi_i|}\left(E - \lambda_{pq}\langle\phi_p|\phi_q\rangle\right) = 0 \\ (|\phi_p\rangle)^+ = \langle\phi_p| \\ \langle\phi_p|\phi_q\rangle = \delta_{pq} \end{array} \right\}$$

$$F|\phi_p\rangle = \sum_q \lambda_{qp}\,|\phi_q\rangle$$

$$F = h + 2\mathcal{J} - \mathcal{K}$$

ρ = ½ density matrix

$$\rho(1,1') = \sum_{i=1}^{N/2} \phi_i(1)\,\phi_i(1')^*$$

Consider unitary transformation

$$(\phi_1' \;\ldots\; \phi_{N/2}') = (\phi_1 \;\ldots\; \phi_{N/2})\,U$$

It mixes occupied orbitals with each other.

Important result:

$$\mathrm{Det}\{\phi_1'\bar{\phi}_1' \ldots\} \equiv \mathrm{Det}\{\phi_1\bar{\phi}_1 \ldots\}$$

$$\rho = \sum \phi_i(1)\phi_i(1')^* \equiv \sum \phi_i'(1)\phi_i'(1')^*$$

so, F is unchanged in form and value by such a transform

$$F\phi_i' = \sum_j U_{ji} F\phi_j$$

$$= \sum_j \sum_q U_{ji} \lambda_{qj} \phi_q$$

$$= \sum_j \sum_q \sum_k U_{ji} \lambda_{qj} (U^{-1})_{kq} \phi_k'$$

$$(U^{-1})_{kq} = U_{qk}^*$$

$$\sum_{qj} U_{qk}^* \lambda_{qj} U_{ji} = (\underline{U}^\dagger \underline{\lambda} \underline{U})_{ki}$$
$$= \underline{\lambda}_{ki}'$$

i.e., $$F\phi_i' = \sum \lambda_{ki}' \phi_k'$$

so ϕ' are also solutions to the Hartree-Fock equations.

An arbitrary (?) but convenient choice are the "canonical orbitals." For some U, λ' is diagonal, so

$$F\phi_i = \varepsilon_i \phi_i$$

is a possible choice.

HARTREE-FOCK EQUATION

Basic Theorems Related to Hartree-Fock Wavefunction

BRILLOUIN THEOREM

Consider unitary mixing of virtual and occupied orbitals.

This does change ψ!

$$\phi_i' \cong \phi_i + \varepsilon\, \phi_a$$

$$\phi_a' \cong \phi_a - \varepsilon\, \phi_i$$

$$\Delta\psi_{ia} \cong \varepsilon \frac{1}{\sqrt{N!}} \operatorname{Det}\{ \ldots(\phi_i\bar{\phi}_a + \phi_a\bar{\phi}_i) \ldots\}$$

$$\Delta E \cong 2 \langle \Delta\psi_{ia}|H|\psi \rangle$$

$$\langle \Delta\psi_{ia}|H|\psi \rangle = 2\varepsilon \langle \phi_i|F|\phi_a \rangle$$

But if $\quad F\phi = \varepsilon\phi$

$$\left\{ \begin{array}{l} \langle \phi_i|F|\phi_a \rangle = \varepsilon_j\, \delta_{ia} \\ \langle \phi_i|\phi_a \rangle = \delta_{ia} \end{array} \right\}$$

so

$$\begin{array}{ccc} \langle \Delta\psi_{ia}|H|\psi \rangle = 0 & \Longleftrightarrow & F_{ia} = 0 \\ \Updownarrow & & \\ \Delta E = 0 & \Longleftrightarrow & E \text{ is stationary} \end{array}$$

Brillouin conditions are frequently used to:

a) derive stationary conditions for more complicated ψ

b) prove perturbation formula:

$$\psi = \psi_{SCF} + \sum C_{i\to a}\,\psi_{i\to a} + \sum_{\substack{i\to a\\ j\to b}} C_{ijab}\,\psi_{ijab} + \cdots$$

$$C_I \cong \frac{\langle \psi_I | H | \psi_{SCF} \rangle}{E_{SCF} - E_I} \qquad \text{first order}$$

where

$C_I = 0$ for single excitations (Brillouin)

$C_I = 0$ for triple, or higher excitations (two-body operator)

Only double excitations contribute to first order.

Molecular One-Electron Properties

$$\langle M \rangle = \int \psi \sum_{i=1}^{N} M(i)\,\psi\,d\tau_1 \ldots d\tau_N$$

SCF:

$$\langle M \rangle = 2\sum_{i=1}^{N/2} \int \psi_i(1)^* M(1)\,\phi_i(1)\,d\tau_1$$

$=$ sum of properties of occupied MO's

Corrections:

$$\psi = \psi_{SCF} + \psi^{(1)}_{double} + \psi^{(2)}_{single} + \psi^{(2)}_{double} + \psi^{(2)}_{triple}$$

$$+ \psi^{(2)}_{quadruple} - \tfrac{1}{2}\, \|\psi^{(1)}\|^2 \;\psi_{SCF}$$

$$\langle \psi_{SCF} | \textstyle\sum M(i) | \psi_L \rangle = 0 \;, \qquad L \geqslant \text{double} \;,$$

so *there is no first order correction to* $\langle M \rangle_{SCF}$!

$$\langle M \rangle \cong \langle M \rangle_{SCF} + 2\langle \psi_{SCF} | M | \psi^{(2)}_{single} \rangle$$

$$+ \langle \psi^{(1)}_{doub} | M | \psi^{(1)}_{doub} \rangle - \|\psi^{(1)}\|^2 \langle M \rangle_{SCF}$$

To a good approximation it is usually true that

$$\langle \psi^1 | M | \psi^1 \rangle \cong \|\psi^{(1)}\|^2 \langle M \rangle_{SCF}$$

so most of $\langle M \rangle^{(2)}$ comes from $\psi^{(2)}_{single}$.

Conclusion: $\langle M \rangle_{SCF}$ should be accurate, but most of correction comes from $\psi^{(2)}_{single}$

$$\psi_{SCF} \xrightarrow[\text{mix}]{} \psi^{(1)}_{double} \xrightarrow[\text{mix}]{} \psi^{(2)}_{single}$$

$\psi^{(2)}_{trip}$ or $\psi^{(2)}_{quad}$ do not matter

KOOPMANS' THEOREMS

Consider positive ion (open shell) in crude approximation of frozen orbitals.

$$\Psi_j^+ = \frac{1}{\sqrt{(N-1)!}} \text{ Det } \{ \phi_1\bar{\phi}_1 \cdots \phi_j\phi_{j+1}\bar{\phi}_{j+1} \cdots \}$$

Approximate ion by

$$\psi^+ \cong \sum C_j \Psi_j^+$$

Do configuration interaction. It requires

$$H_{ij} = \langle \Psi_i^+ | H | \psi_j^+ \rangle$$

$$\langle \Psi_i^+ | H | \Psi_j^+ \rangle = -\langle \phi_j | F | \phi_i \rangle + E^o \delta_{ij}$$

$$= (E^o - \varepsilon_i)\delta_{ij} \quad \begin{pmatrix}\text{canonical} \\ \text{orbitals}\end{pmatrix}$$

so $\underline{H}$ is diagonal and

$$\boxed{I_i = -\varepsilon_i}$$

for canonical orbitals. Similarly for negative ion:

$$\psi_a^- = \frac{1}{\sqrt{(N+1)!}} \text{ Det } \{ \phi_1 \cdots \bar{\phi}_{N/2}\phi_a \}$$

$$\langle \psi_a^- | H | \psi_b^- \rangle = \left\{ \begin{matrix} 0 & a \neq b \\ E^o + \varepsilon_a & a = b \end{matrix} \right\}$$

for canonical orbitals

$$\boxed{EA \cong -\varepsilon_a}$$

<u>Errors</u>

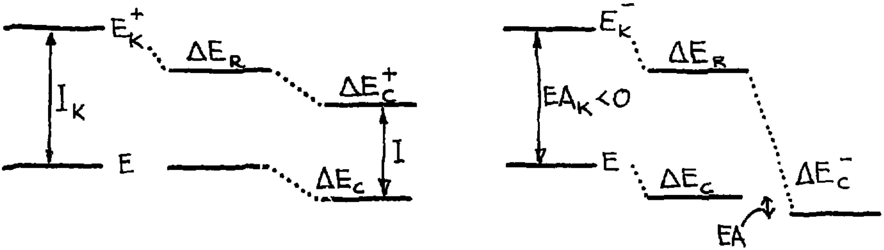
ΔE_c = difference in correlation E

E_c = "correlation E"

$\equiv$ "error in E_{SCF}"

ΔE_R = relaxation E

$I \cong I_K \quad , \quad EA \ncong EA_K$



$$E_{SCF} = \sum_i 2h_{ii} + \sum_{i,j} (2J_{ij} - K_{ij}) + V_{NN}$$

$$E_{SCF} = \sum 2\varepsilon_i - V_{ee} + V_{NN}$$

ε_i includes J_{ij}

ε_j includes J_{ij}

$\varepsilon_i + \varepsilon_j$ includes J_{ij} twice

<u>or</u>

$$E_{SCF} = \sum (\varepsilon_i + h_{ii}) + V_{NN} .$$

NOTE $$E_{SCF} \neq \sum 2\varepsilon_i .$$

$$\frac{\delta E}{\delta R} \overset{??}{\neq} \frac{\delta \varepsilon_{HOMO}}{\delta R}$$

as often assumed. Very difficult to justify!

<u>ROOTHAAN EQUATIONS</u>

Expand $$\phi_i = \sum x_{ai} f_a$$

where x coefficients

f_a "atomic orbitals"

"basis functions"

same set of f_a used for all MO's, LCAO-MO-SCF.

Determine x to minimize $\langle H \rangle$ for fixed f. (Perhaps also optimize f?) Following previous derivation taking $\delta\phi \leftrightarrow \delta x$ gives matrix equation

$$\langle f_a|F\phi_i\rangle = \varepsilon_i\langle f_a|\phi_i\rangle$$

Expanding ϕ_i

$$\sum_b \langle f_a|F|f_b\rangle\, x_{bi} = \varepsilon_i \sum_b \langle f_a|f_b\rangle\, x_{bi}$$

where

$\langle f_a|f_b\rangle = S_{ab}$ overlap matrix

$\langle f_a|F|f_b\rangle = F_{ab}$ Fock matrix

$$\underset{\sim}{F}\,\underline{x} = \underline{S}\,\underline{x}\,\underline{\varepsilon}$$

$$F_{ab} = h_{ab} + 2\mathcal{J}_{ab} - \mathcal{K}_{ab}$$

$$\rho = \sum_i \phi_i\phi_i^* \equiv \sum_{c,d} P_{c,d}\, f_c f_d^*$$

$$P_{cd} = \sum_i x_{ic}x_{id} = \tfrac{1}{2} \text{ bond order;}$$

$= \text{charge density matrix}$

$$\mathcal{J}_{ab} = \sum_{cd} P_{cd}\,[ab\|cd]$$

$$\mathcal{K}_{ab} = \sum_{cd} P_{cd}\,[ac\|bd]$$

$$[ab\|cd] = \int f_a^{(1)} f_b^{(1)} \frac{1}{r_{12}} f_c^{(2)} f_d^{(2)}\, d\tau_1\, d\tau_2$$

Solve iteratively

$$\begin{tikzcd} \text{guess } \underline{x} \text{ or } \underline{P} \arrow[d] \\ \text{construct } F \arrow[d] \\ \text{solve } \underline{F}\underline{Y} = \underline{S}\underline{Y}\underline{\varepsilon} \arrow[uu, bend right, "\text{iterate}"'] \end{tikzcd}$$

often $x_{next} \leftarrow y$

sometimes $$P_{next} = \frac{P_{old} + \alpha P(y)}{1 + \alpha}$$

sometimes diagonal of F is modified to improve convergence of $y \rightarrow x_{SCF}$. At self-consistency

$$y_{out} = x_{in}$$

Interpretation of ψ is usually done through

$$\underset{\sim}{B} = 2\underset{\sim}{P}$$

$$2 \sum \phi_i \phi_i^* = \sum B_{ab} f_a f_b^*$$

$$\langle M \rangle = \sum B_{ab} M_{ba}$$

$$\langle 1 \rangle = N = \sum B_{ab} S_{ba}$$

$$E = \sum B_{ab} (F_{ba} + h_{ba}) + V_{NN}$$

MULLIKEN POPULATION ANALYSIS

Assume f_a are literally "atomic orbitals." Then

$$\sum B_{ab} S_{ba}$$

can be partitioned

$$S_{ba} = \begin{cases} 1 & a=b \\ 0 & a \neq b \end{cases} \quad \text{a and b on same atom}$$

Define

$$q_a = B_{aa} + \sum_{B(\neq A)} \sum_{\substack{b \\ \text{on } B}} B_{ab} S_{ba}, \quad \text{for a "on" A.}$$

$B_{ab} S_{ba}$ "overlap population" generally proportional to bond strength. B_{ab} "bond order" also proportional to bond strength for S at R_e.

Difficulties:

- q sometimes > 2
- q sometimes < 0
- q arbitrary partitioning of overlap population between a and b equally
- q difficult to extend to other basis sets

Advantages:

$$\sum_{i \text{ on } A} q_i = q_A .$$

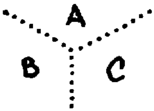
independent of molecular rotation even if basis does not rotate.

OTHER POPULATION SCHEMES

1. Sphere charge.

2. Proportionate splitting of overlap population.

3. Atomic boundary:



4. Extended basis sets:

   – one center expansion

   – overcomplete multicenter set

   – how to handle?

     Project result onto minimum "atomic orbital" set or onto "scaled AO set"

5. Non-orthogonal sets on an atom, e.g., STO, Gaussians, etc. Partitioning of "one-center" overlap equally can give strange results.

BASIS SETS IN COMMON USE

Slater type orbital (atoms, diatomics)

$$r^{n-1}\, e^{-\zeta r}\, Y_{LM}(\theta,\phi)$$

Slater type orbital (polyatomics)

$$r^{n-\ell-1}\, e^{-\zeta r}\, \mathcal{P}_{L,M}(x,y,z)$$

$\mathcal{P}_{L,M}$ = real spherical polynomial

Atomic orbital — literally solution to atomic SCF equations. May be of Roothaan type, i.e., expanded in a one-center basis of STO's or Gaussians

Gaussian lobe $e^{-\alpha r_A^2}$

Cartesian Gaussian $x_A^n\, y_A^\ell\, z_A^m\; e^{-\alpha r_A^2}$

Contracted Gaussians

$$f_a = \sum_\mu C_{\mu a}\, g_\mu \quad , \quad \text{fixed } C_{\mu a}$$

<u>STO BASIS SET NAMING</u>

Minimum: One STO for each occupied AO.

Double zeta: Two STO for each occupied AO with different ζ.

Split valence: One STO for each core AO, two STO for each valence AO.

Polarization: One set of STO's of higher L than any occupied in atoms.

Common level of "accurate" calculations:
double zeta + polarization.

Rydberg orbitals: Approximations of diffuse orbitals used in excited states of atoms

COMMONLY USED CONTRACTIONS

Pople

$\underbrace{\text{STO 3G}}_{\text{minimum}}$ $\qquad$ $\underbrace{\text{431G} \quad \text{321G}}_{\text{split valence}}$ $\qquad$ $\underbrace{\text{431G}^*}_{\text{split valence + polarization}}$

Dunning/Huzinaga

(9s/5p) $\longrightarrow$ [4s/2p] DZ

(9s/5p) $\longrightarrow$ [3s/2p] split valence

H (4s) $\longrightarrow$ [2s]

Raffenetti extended; not disjoint

[5s/3p/1d]

Duijneveldt/McMurchie

(14s/8p) $\longrightarrow$ [8s/6p] + 2d

H (8s) $\longrightarrow$ [6s]

Even tempered

$e^{-\alpha\beta^n r^2}$, $n = 0,1,2...$

OPEN SHELL HIGH SPIN HARTREE-FOCK

$$\psi = \frac{1}{\sqrt{N!}} \left\{ \phi_1 \bar{\phi}_1 \phi_2 \bar{\phi}_2 \cdots \phi_k \bar{\phi}_k \phi_{k+1} \phi_{k+2} \right\}$$

Unrestricted Hartree-Fock

$$\psi = \frac{1}{\sqrt{N!}} \left\{ \phi_1 \phi_2 \cdots \phi_p \, \bar{\phi}_1' \bar{\phi}_2' \cdots \bar{\phi}_q' \right\}$$

No longer S^2 eigenfunction

$\langle \phi_i \phi_j' \rangle = 0$ by spin orthogonality

$$\frac{\delta}{\delta \langle \phi_i |} \left\{ E - \sum \lambda_{ij} \langle \phi_i | \phi_j \rangle - \sum \bar{\lambda}_{ij} \langle \bar{\phi}_i | \bar{\phi}_j \rangle \right\} = 0$$

$$F_\alpha \phi_i = \varepsilon_i \phi_i$$

$$F_\beta \phi_i = \varepsilon_i' \phi_i'$$

canonical choice of orbitals

$$F_\alpha = h + \mathcal{J}_\alpha + \mathcal{J}_\beta - \mathcal{K}_\alpha$$

$$F_\beta = h + \mathcal{J}_\alpha + \mathcal{J}_\beta - \mathcal{K}_\beta$$

$$\rho_\alpha = \sum \phi_i \phi_i^*$$

$$\rho_\beta = \sum \phi_i' \phi_i'^*$$

In Roothaan form

$$\underset{\sim}{F}_\alpha \underset{\sim}{x}_i = \varepsilon_i \underset{\sim}{S} \underset{\sim}{x}_i$$

$$\underset{\sim}{F}_\beta \underset{\sim}{x}_i = \varepsilon_i \underset{\sim}{S} \underset{\sim}{x}_i$$

Best single determinant with orbitals of pure spin (α or β)

Advantages of UHF:

- correct dissociation
- simplicity
- more general Brillouin theorem and Koopmans' theorem

Disadvantages:

- not S^2 eigenfunction

EXAMPLE: H_2 DZP Basis

RHF

$$\psi_{SCF} = \frac{1}{\sqrt{2}} \text{ Det } \{\phi_1 \bar{\phi}_1\}$$

$$\phi_1 = c_1(1s_A + 1s_B) + c_2(1s_A' + 1s_B') + c_3(2p_{ZA} + 2p_{ZB})$$

Large R: ψ is 50% ionic

$$\lim R \to \infty \ \psi_{SCF} \text{ is not } \psi$$

$$\psi \to (1s_A 1s_B + 1s_B 1s_A)$$

but $c_2 \not\to 0$

UHF, there is critical R^*

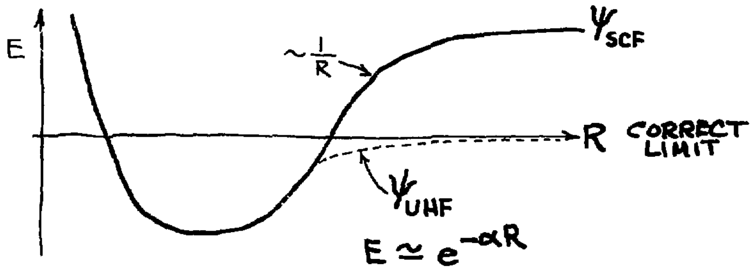
for $R < R^*$ $\qquad \psi_{UHF} \equiv \psi_{RHF}$

for $R > R^*$ $\qquad \psi_{UHF} = \frac{1}{\sqrt{2}} \text{Det} \{\phi_1 \bar{\phi}_1\}$

$\phi_1 \to 1s_A$

$c_2 \to 1s_B$

$\psi \to 1s_A \overline{1s_B}$



Notice neither curve has R^{-6} shape. UHF curve has discontinuous slope at R^*. UHF is only weakly bonding. ψ_{UHF} is not ψ_{VB} so it does not give strong valence bond

$$\psi_{UHF} \rightarrow 50\% \ (S=0) + 50\% \ (S=1)$$

$$\psi_{SCF} \rightarrow 50\% \ \psi_{VB} + 50\% \ \psi_{ionic}$$

Both wrong!

In general, there is difficulty with UHF when two states *of different S are close in energy.*

SPIN DENSITIES

$$\rho_s = \frac{\rho_\alpha - \rho_\beta}{N_\alpha - N_\beta}$$

HYPERFINE SPLITTING PARAMETERS

$$\bar{a}_A = g_e\beta_e \ g_N\beta_N \ \frac{8\pi}{3} \ \rho_s(A)$$

$$a_{zz} = \bar{a}_A + g_e\beta_e \ g_N\beta_N \int \rho_s \ \frac{3z_A^2 - r_A^2}{r_A^5} \ d\tau$$

SPIN POPULATIONS

$$\rho_\alpha - \rho_\beta = \sum_{a,b} P^s_{ab} \ f_a f_b$$

$$q_a^{(s)} = \text{spin population}$$

For σ radicals: "a" is usually OK (±10%)

For π radicals: "a" is zero for RHF

"a" is non-zero *but* very inaccurate for UHF

$$H - \dot{C} \begin{array}{l} / \text{ proton hyperfine} \\ \backslash\ a_H \sim (-27 \text{ gauss}) \left(q^{(s)}_{\pi_C}\right) \end{array}$$

SPIN PROJECTION

$$\prod_{S \neq S'} \frac{\mathcal{S}^2 - S'(S'+1)}{S(S+1) - S'(S'+1)} \psi_{UHF}$$

produces spin eigenfunction but *not* better spin distribution (usually).

LOCALIZED ORBITALS

Recall ψ_{SCF} is unchanged by unitary transformation among occupied orbitals.

H_2 : $$\psi = \frac{1}{\sqrt{2}} \{\phi_1 \bar{\phi}_1\}$$

He_2 : $$\psi = \frac{1}{\sqrt{4!}} \{\phi_1 \bar{\phi}_1 \ \phi_2 \bar{\phi}_2\}$$

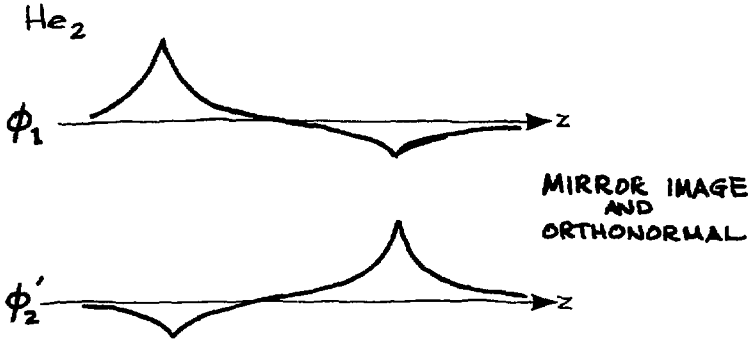
where

$$\phi_1 \cong (1s_A + 1s_B)/\sqrt{2(1+S)}$$

$$\phi_2 \cong (1s_A - 1s_B)/\sqrt{2(1-S)}$$

$$\downarrow$$

$$\frac{1}{\sqrt{2}}(\phi_1+\phi_2) = \phi_1' = \frac{1s_A}{2}\left(\frac{1}{\sqrt{1+S}} + \frac{1}{\sqrt{1-S}}\right) + \frac{1s_B}{2}\left(\frac{1}{\sqrt{1+S}} - \frac{1}{\sqrt{1-S}}\right)$$



$$\psi \equiv \frac{1}{\sqrt{4!}}\{\phi_1'\bar{\phi}_1'\ \phi_2'\bar{\phi}_2'\}$$

NOTE: Equivalence transformation

$$(\phi_1' \cdots \phi_{N/2}') = (\phi_1 \cdots \phi_{N/2})\, W$$

$|W| \neq 0$ also leaves ψ unchanged

$$\psi \equiv \frac{1}{\sqrt{N!}} |W|^{-1} \{\phi_1'\bar{\phi}_1' \cdots \phi_{N/2}'\bar{\phi}_{N/2}'\}$$

$$\Psi \equiv \frac{1}{\sqrt{4!}} \left\{\overline{1s}_A 1s_A \; 1s_B \overline{1s}_B\right\} / (1 - S^2)$$

Non-orthogonal basis fully localized.

The formula for E and ρ in non-orthogonal basis is complicated. Most authors define best orthogonal localized orbitals.

RUEDENBERG:

maximize $\sum_i J_{ii}$

minimize $\sum_{i \neq j} (2J_{ij} - K_{ij})$

OTHER POSSIBILITIES:

minimize $\sum_i (\sigma_x^2 + \sigma_y^2 + \sigma_z^2)_i$

minimize $\sum_{i \neq j} \int |\phi_i|^2 \; |\phi_i|^2 \, d\tau$

minimize $\sum_i \frac{1}{\int \phi_i^4 \, d\tau} = \sum_i$ orbital volume

DISADVANTAGES OF LOCALIZED ORBITAL

- Broken symmetry (CI cannot take advantage of symmetry)
- Banana bonds
- Non-negligible tails

ADVANTAGES

- CI more compact?
- Transferable?

$$C_{100}H_{202} \rightarrow 2 \cdot C_{50}H_{51}$$

{ Do all MO's change or only a few? }

Conceptual!

ACTUAL PRACTICE

Seldom used except conceptually.

No good for spectra or ionization?

Or are they better?

ITERATIVE METHODS FOR SOLVING THE HARTREE-FOCK EQUATION

Try to solve $\qquad \underset{\sim}{x} = \underset{\sim}{G}(\underset{\sim}{x})$

by procedure $\qquad \underset{\sim}{y} = \underset{\sim}{G}(\underset{\sim}{z})$

Convergence?

$$G_i(\underset{\sim}{z}) \cong G_i(\underset{\sim}{x}) + \sum_j \left(\frac{\partial G_i}{\partial G_j}\right)_x (z_j - x_j)$$

$$(y_i - x_i) \cong \sum_j \left(\frac{\partial G_i}{\partial x_j}\right)_x (z_j - x_j)$$

$$\|y - x\| \leq \left\|\frac{\partial \underset{\sim}{G}}{\partial \underset{\sim}{x}}\right\|_{\underset{\sim}{x}} \cdot \|z - x\|$$

Converges for $\qquad \left\|\frac{\partial \underset{\sim}{G}}{\partial \underset{\sim}{x}}\right\| < 1$

May diverge if $\qquad \left\|\frac{\partial \underset{\sim}{G}}{\partial \underset{\sim}{x}}\right\| > 1$

Second order process

$$\Longrightarrow \left.\frac{\partial G_i}{\partial x_j}\right|_{\underset{\sim}{x}} = 0, \qquad \left\|\frac{\partial \underset{\sim}{G}}{\partial \underset{\sim}{x}}\right\| = 0$$

FOCK/BRILLOUIN/CI PROCESSES

$\underset{\sim}{z}$ = current guess to MO coefficients

$\underset{\sim}{y}$ = new guess

$\underset{\sim}{x}$ = correct coefficients

$$\underset{\sim}{Y} = \underset{\sim}{Z}\,\underset{\sim}{U} \Longleftrightarrow \underset{\sim}{U} = \underset{\sim}{Z}^{-1}\,\underset{\sim}{Y}$$

$\underset{\sim}{U}$ = unitary matrix

(${\underset{\sim}{z}}$ and ${\underset{\sim}{y}}$ both orthonormal sets)

FOCK PROCEDURE

$$\underset{\sim}{F}\big(\rho(z)\big)\,\underset{\sim}{Y} = \underset{\sim}{S}\,\underset{\sim}{Y}\,\underset{\sim}{\varepsilon}$$

$$\underbrace{Z^{\dagger}\,F\,Z}_{\tilde{F}}\;\underbrace{Z^{-1}\,Y}_{U} = \underbrace{Z^{\dagger}\,S\,Z}_{1}\;\underbrace{Z^{-1}\,Y}_{U}\,\varepsilon$$

i.e.,

$$\underset{\sim}{\tilde{F}}\,\underset{\sim}{U} = \underset{\sim}{U}\,\underset{\sim}{\varepsilon}$$

To first order, find U by perturbation theory

(near convergence, $\tilde{F}$ almost diagonal)

$$\underbrace{\phi_j'}_{\text{"y"}} = \sum U_{ij}\,\underbrace{\phi_i}_{\text{"z"}}$$

$$U_{ij} \cong \frac{\tilde{F}_{ij}}{\tilde{F}_{jj} - \tilde{F}_{ii}} \qquad (i \neq j)$$

<u>GUEST-SAUNDERS</u> (ideas from level shift paper)

True second order

$$U_{ij} = - \underbrace{\left[\frac{\partial^2 E}{\partial U_{ij} \; \partial U_{k\ell}} \right]^{-1}}_{\substack{\text{super} \\ \text{matrix}}} \cdot \underbrace{\left[\frac{\partial E}{\partial U_{k\ell}} \right]}_{\substack{\text{super} \\ \text{vector}}}$$

Approximate second order

$$U_{ij} \cong - \left(\frac{\partial^2 E}{\partial U_{ij}^2} \right)^{-1} \left(\frac{\partial E}{\partial U_{ij}} \right)$$

Evaluating derivatives gives

$$U_{ij} \cong \frac{\tilde{F}_{ij}}{\tilde{F}_{jj} - \tilde{F}_{ii} + J_{ij} - 3K_{ij}} \qquad (i \neq j)$$

INO/CI/SCF

$$\psi \cong \psi_{(z)} + \sum_{j,i} C_{i\leftarrow j} \, \psi_{i\leftarrow j}$$

Determine $C_{i\leftarrow j}$ by CI. Compare

$$\psi \cong \psi_{(z)} + \sqrt{2} \sum_{j,i} U_{ij} \, \psi_{i\leftarrow j} + O(u^2)$$

so to first order

$$U_{ij} \cong C_{i\leftarrow j}/\sqrt{2}$$

PERTURBATION THEORY FOR CI EIGENVECTOR

$$C_{i \leftarrow j} \cong \frac{\langle \psi_{i \leftarrow j} | H | \psi_{(z)} \rangle}{\langle \psi_z | H | \psi_z \rangle - \langle \psi_{i \leftarrow j} | H | \psi_{i \leftarrow j} \rangle}$$

Evaluating matrix elements gives

$$U_{ij} \cong \frac{\tilde{F}_{ij}}{\tilde{F}_{jj} - \tilde{F}_{ii} + J_{ij} - 3K_{ij}}$$

to first order.

CONCLUSION

Fock iteration is sensible, IF

$$\frac{\partial^2 E}{\partial U_{ij}\ \partial U_{k\ell}}$$

is diagonal dominant, and if

$$\tilde{F}_{jj} - \tilde{F}_{ii}$$

has same sign as $\tilde{F}_{jj} - \tilde{F}_{ii} + J_{ij} - 3K_{ij}$

(j occ., i unocc.)

Level Shift: add constant to $\tilde{F}_{ii}$ (i unocc) to make

$\tilde{F}_{jj} - (\tilde{F}_{ii} + \alpha)$ approximate

$\tilde{F}_{jj} - (\tilde{F}_{ii} - J_{ij} + 3K_{ij})$

or to make $\|U-1\|$ small.

SEMIEMPIRICAL SCF THEORY

Lecture 4

by

Michael Zerner

Guelph-Waterloo Centre for Graduate Work in Chemistry
University of Guelph
Guelph, Ontario, Canada

INTEGRAL APPROXIMATIONS

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\sigma,\lambda} P_{\sigma\lambda} \underbrace{(\mu\nu|\sigma\lambda)}_{\substack{N^4 \\ \text{integrals}}} - \sum P^{\alpha}_{\sigma\lambda} (\mu\sigma|\nu\lambda)$$

This is such a large problem that HF theory is geared to integral evaluation and processing. The fastest method for SCF is probably the "super matrix" formalism:

$$F^{\alpha}_{\beta\,\mu\nu} = h_{\mu\nu} + \sum_{\sigma,\lambda} P_{\sigma\lambda}\, \mathcal{P}_{\mu\nu\sigma\lambda} \mp \tfrac{1}{2} \sum \mathcal{K}_{\mu\nu\sigma\lambda}\, \rho_{\sigma\lambda}$$

where

$$\mathcal{P}_{\mu\nu\sigma\lambda} = \text{supermatrix element}$$

$$= (\mu\nu|\sigma\lambda) - \tfrac{1}{2} \mathcal{K}_{\mu\nu\sigma\lambda}$$

$$\mathcal{K}_{\mu\nu\sigma\lambda} = \tfrac{1}{2} \left[(\mu\sigma|\nu\lambda) + (\mu\lambda|\nu\sigma) \right]$$

$\rho_{\sigma\lambda}$ is spin density $= P^{\alpha}_{\sigma\lambda} - P^{\beta}_{\sigma\lambda}$

$$P_{\sigma\lambda} = P^{\alpha}_{\sigma\lambda} + P^{\beta}_{\sigma\lambda}$$

REDUCING THE NUMBER OF INTEGRALS

Reduction Based on Integral Size

$$(\mu\sigma|\nu\lambda) \sim \frac{\Delta_{\bar{\mu}\bar{\sigma}} \Delta_{\bar{\nu}\bar{\lambda}}}{4} \left[(\mu\mu|\nu\nu) + (\mu\mu|\lambda\lambda) + (\sigma\sigma|\nu\nu) + (\sigma\sigma|\lambda\lambda) \right]$$

where

$$\Delta_{\bar{\mu}\bar{\sigma}} = (\bar{\mu}|\bar{\sigma})$$

This is not so important for small systems.

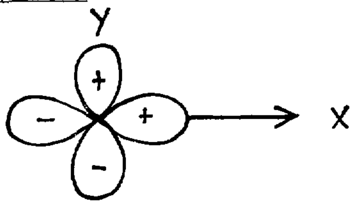
VRDDO = Variable retention of differential diatomic overlap by Popkie and Kaufman

VRDDO + core potential = VRDDO - MODPOT

Retention by Systematic Approximation or Neglect

Systematic approximation or neglect involves both symmetry and balance.

Symmetry:



$$\bigcirc\ \chi_\sigma^*(2)\,\chi_\sigma(2)$$

$$(xy|\sigma\sigma) = 0 \qquad \text{by ZDO}$$

$$xy = \underbrace{P_x(1)\,P_y(1)}_{\text{ADO}}$$

$$x \rightarrow \frac{1}{\sqrt{2}}(x+y)$$

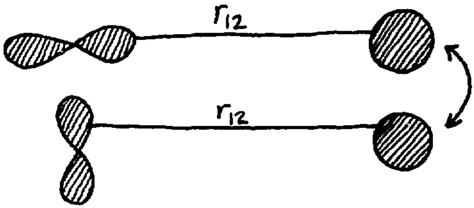
$$y \rightarrow \frac{1}{\sqrt{2}}(x-y)$$

$$0 = (xy|\sigma\sigma) \rightarrow \tfrac{1}{2}\{-(xx|\sigma\sigma) + (yy|\sigma\sigma)\}$$

<u>so</u>

$$(xx|\sigma\sigma) \approx (yy|\sigma\sigma) \qquad \text{to have ZDO}$$

$$(xx|\sigma\sigma) \approx (yy|\sigma\sigma) = (\bar{p}\bar{p}|\sigma\sigma)$$



these interactions are considered the same

In INDO (intermediate neglect of differential overlap):

$$(yy|yy) = F_o + \frac{4}{25} F^2$$

$$(xx|yy) = F_o - \frac{2}{25} F^2$$

We must have

$$(xy|xy) = \frac{3}{25} F^2$$

for s,p basis for all integrals of the form $(ii|jj)$ or $(ij|ij)$ that we considered for atoms.

Note that Δ is OK, so

$$\beta_{\nu\mu} = \Delta_{\mu\nu}\beta^{o}_{\mu\nu} \qquad \text{is OK}$$

$$\beta_{\nu\mu} \cong \Delta_{\mu\nu}\beta^{o}_{\mu\nu} + \Delta^{2}_{\mu\nu}\beta'_{\mu\nu} \qquad \text{is not OK}$$

We will return to this later.

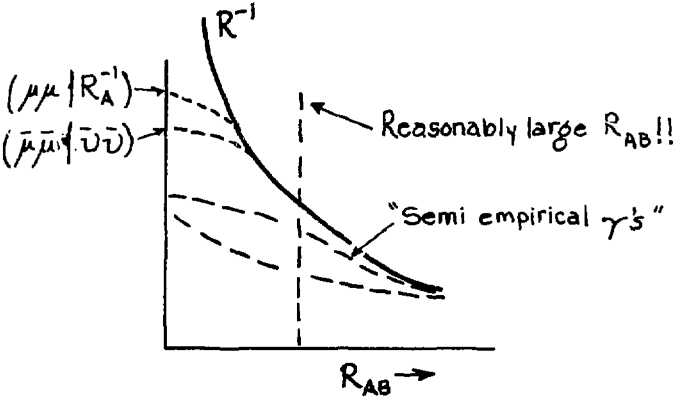
Balance

Consider two idential atoms

Their electrostatic energy is

$$E_{es} \approx \frac{Z_A \; Z_B}{R_{AB}} + Z_A Z_B \; \gamma_{\bar{A}\bar{B}} - 2 Z_A Z_B (\bar{A}\bar{A} | R_B^{-1})$$

$$= \quad 0 \qquad \text{at reasonably large } R_{AB} \quad .$$



Usually for balance in Fock matrix

$$(\bar{\mu}\bar{\mu} | R_A^{-1}) + V(\text{core}) + V(\text{orthog}) \;\; = \;\; (\bar{\mu}\bar{\mu} | \bar{\nu}\bar{\nu}) \quad .$$

set $\qquad (\bar{\mu}\bar{\mu} | R_A^{-1}) \;\; = \;\; (\bar{\mu}\bar{\mu} | \bar{\nu}\bar{\nu}) \quad .$

If $\gamma_{AB} = (\bar{A}\bar{A}|\bar{B}\bar{B})$ from semi-empirical approximations, then we must scale nuclear repulsion energy. Spectroscopic INDO and CNDO theories at present do not give geometries!

Balance: a more subtle example!

Consider Mulliken population

$$N = \sum_i n_i = \sum_i n_i \langle \phi_i | \phi_i \rangle$$

where

$$N = \text{number of electrons}$$

$$n_i = 0 \text{ or } 1 = \text{occupation of } \phi_i$$

$$N = \sum_{i,\mu,\nu} n_i \, C_{i\mu} C_{i\nu} \Delta_{\mu\nu}$$

$$= \sum_{\mu,\nu} P_{\mu\nu} \Delta_{\mu\nu} = \sum_{\mu} \left\{ \sum_{\nu} P_{\mu\nu} \Delta_{\mu\nu} \right\}$$

$$M_{\mu\mu} = \sum_{\nu} P_{\mu\nu} \Delta_{\mu\nu}$$

$$= \text{Mulliken Orbital Population}$$

$$M_A = \sum_{\mu}^{A} M_{\mu\mu}$$

Now

$$J_{\mu\nu} = \sum_{\sigma,\lambda} P_{\sigma\lambda}(\mu\nu|\sigma\lambda)$$

$$\{\chi_\mu\} \rightarrow \{x'_\mu\}$$

$$J_{\mu\nu} = \sum_{\sigma,\lambda} P'_{\sigma\lambda}(\mu\nu|\sigma'\lambda')$$

This is exact. Suppose we choose $\{\chi_\mu\} \ni (\mu\nu|\sigma'\lambda')$ very small for $\sigma' \neq \lambda'$. Then

$$J_{\mu\nu} = \sum_{\sigma,\lambda} P'_{\sigma\lambda}(\mu\nu|\sigma'\lambda') = \sum_{\sigma} P'_{\sigma\sigma}(\mu\nu|\sigma'\sigma')$$

The four center integrals can thus be dropped in a systematic rational fashion, so $N^4 \rightarrow N^3$.

What about three center integrals?

$$F^{AB}_{\mu\nu} = (\mu|t|\nu) - \sum_B (\mu|R_B^{-1}|\nu)Z_B$$

$$+ \sum P'_{\sigma\sigma}(\mu\nu|\sigma'\sigma') - K_{\mu\nu}$$

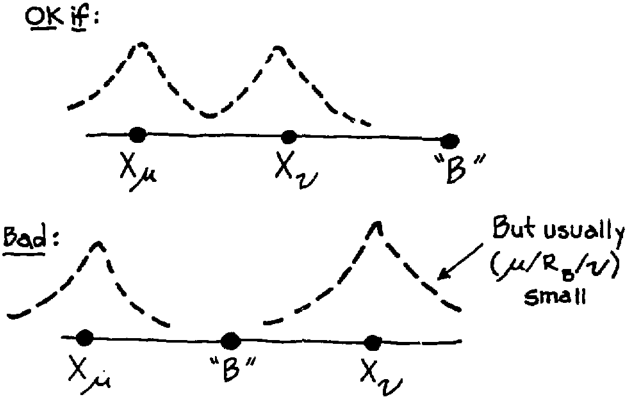
The three-center terms

$$\approx \sum_B \left\{ P'_B(\mu\nu|\bar{\sigma}'\bar{\sigma}') - Z_B(\mu|R_B^{-1}|\nu) \right\}$$

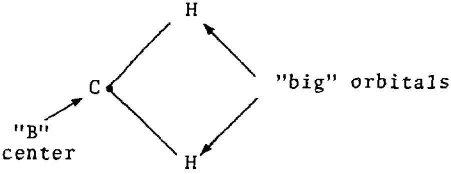
Drop all three center terms

$$\sum_B Q_B(\mu|R_B^{-1}|\nu)$$

where $Q_B = P'_B - Z_B$ *if* Q_B *is small*, and $(\bar{\mu}|R_B^{-1}|\nu) \approx (\mu\nu|\bar{\sigma}'\bar{\sigma}')$.



Three center terms of the type

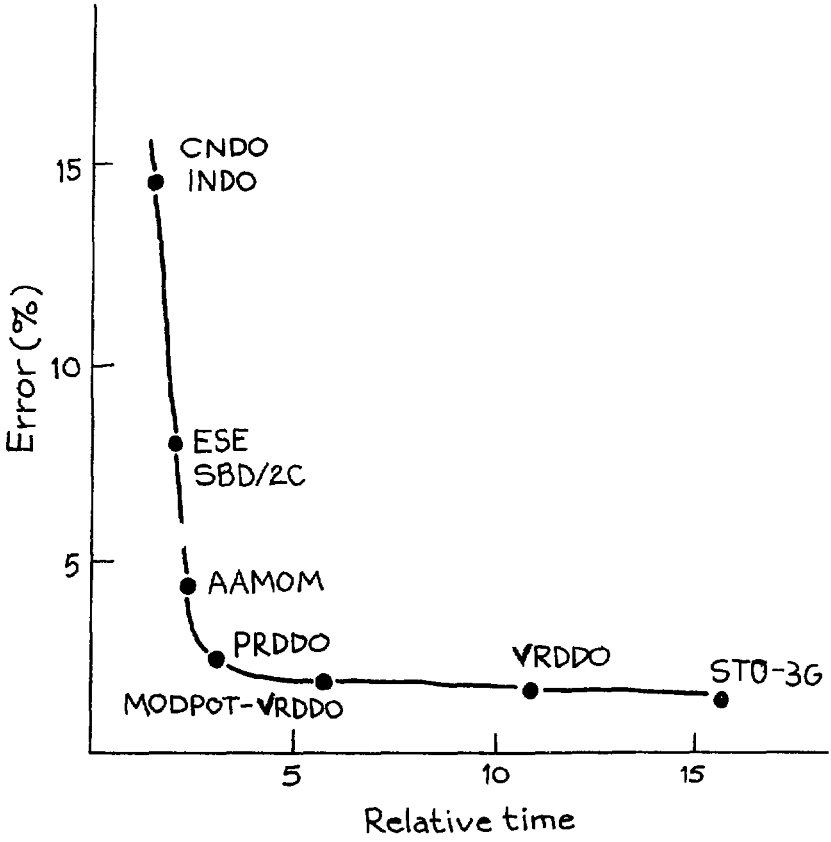


are treated differently by certain methods.

PRDDO and AAMOM methoa keep these terms.

PRDDO = partial retention of differential diatomic overlap

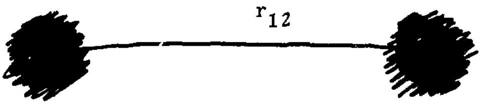
AAMOM = an approximate MO method



from Halgren, Libscomb, et al.
JACS 100, 6595 (1978)

N^2 ZDO Methods

$$F^{\alpha}_{\mu\mu} = U_{\mu\mu} + \sum [P_{\sigma\lambda}(\mu\mu|\sigma\lambda) - P^{\alpha}_{\sigma\lambda}(\mu\sigma|\mu\lambda)]$$



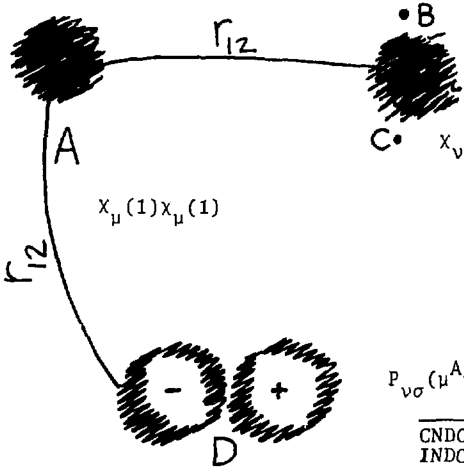
$\overset{AA\ BB}{P_{\nu\nu}(\mu\mu|\nu\nu)}$

$\chi_\mu(1)\chi_\mu(1)$ $\chi^*_\nu(2)\chi_\nu(2)$

CNDO = *Complete neglect of DO*

INDO = Intermediate neglect of DO

NDDO = Neglect of differential diatomic O



$P_{\nu\sigma}(\mu^A\mu^A|\nu^B\sigma^C)$

$\chi_\nu(2)\chi_\sigma(2) \xrightarrow[\substack{CNDO\\INDO\\NDDO}]{} 0$

$\chi_\mu(1)\chi_\mu(1)$

$P_{\nu\sigma}(\mu^A\mu^A|\nu^B\sigma^C)$

$\xrightarrow[\substack{CNDO\\INDO}]{} 0$

$F^K,\ G^K\ ,\qquad K > 0 \xrightarrow[CNDO]{} 0$

ZDO Integral Approximations

1) NDDO

$$\chi_\mu^A(1)\ \chi_\nu^B(1)\ d\tau(1) \quad = \quad \delta_{AB}\ \chi_\mu^A(1)\chi_\mu^A(1)\ d\tau(1)$$

2) CNDO

$$\chi_\mu(1)\chi_\mu(1)\ d\tau(1) \quad = \quad \delta_{\mu\nu}\ \chi_\mu(1)\,\chi_\nu(1)\ d\tau(1)$$

3) *INDO = CNDO + all one centre terms*

$$(\mu^A\nu^B|\sigma^c\lambda^d) \quad = \quad \delta_{AB}\delta_{CD}\delta_{AC}(\mu^A\nu^A|\sigma^A\lambda^A)$$
$$(1 - \delta_{AC})\,\delta_{\mu\nu}\delta_{\sigma\lambda}(\mu^A\mu^B|\sigma^C\sigma^C)$$

Diagonal: $\mu \in A$

Full

$$F_{\mu\mu}^\alpha \quad = \quad U_{\mu\mu} + \sum P_{\sigma\lambda}(\mu\mu|\sigma\lambda) \ - \ P_{\sigma\lambda}^\alpha(\mu\sigma|\mu\lambda) \ - \ \sum_{B\neq A}(\mu\mu|R_B^{-1})Z_B$$

NDDO

$$F_{\mu\mu}^\alpha \quad = \quad U_{\mu\mu} + \sum_{[\sigma,\lambda]\varepsilon B} P_{\sigma\lambda}(\mu\mu|\sigma\lambda) - \sum_{[\sigma,\lambda]\varepsilon A} P_{\sigma\lambda}^{\alpha}(\mu\sigma|\mu\lambda) - \sum_{B\neq A} Z_B(\mu\mu|R_B^{-1})$$

INDO

$$F_{\mu\mu}^\alpha \quad = \quad U_{\mu\mu} + \sum_{[\sigma,\lambda]\varepsilon A} P_{\sigma\lambda}(\mu\mu|\sigma\lambda) \ - \ P_{\sigma\lambda}(\mu\sigma|\mu\lambda)$$
$$+ \sum_{\sigma\neq A} P_{\sigma\sigma}(\bar{\mu}\bar{\mu}|\bar{\sigma}\bar{\sigma}) \ - \ \sum_{B\neq A} Z_B(\bar{\mu}\bar{\mu}|R_B^{-1})$$

for H and first row $\sigma=\lambda$.

CNDO

$$F^{\alpha}_{\mu\mu} = U_{\mu\mu} + \sum_{\sigma} P_{\sigma\sigma}(\bar{\mu}\bar{\mu}|\bar{\sigma}\bar{\sigma}) - P^{\alpha}_{\mu\mu}(\bar{\mu}\bar{\mu}|\bar{\mu}\bar{\mu}) - \sum_{B\neq A} Z_B(\bar{\mu}\bar{\mu}|R_B^{-1}$$

Off-Diagonal $\{\mu,\nu\} \in A$

NDDO

$$F^{\alpha}_{\mu\nu} = \overbrace{t_{\mu\nu} - \sum_{B} (\mu\nu|R_B^{-1}) Z_B}^{\equiv\, \beta_{\mu\nu}} + \sum_{[\sigma,\lambda]\in A} P_{\sigma\lambda}(\mu\nu|\sigma\lambda)$$

$$- \sum_{[\sigma,\lambda]\in A} P^{\alpha}_{\sigma\lambda}(\mu\sigma|\nu\lambda)$$

For H and first row

$$F^{\alpha}_{\mu\nu} = \beta_{\mu\nu} + \sum_{[\sigma,\lambda]\in B} P_{\sigma\lambda}(\mu\nu|\sigma\lambda) - P^{\alpha}_{\mu\nu}[(\mu\mu|\nu\nu) + (\mu\nu|\mu\nu)]$$

INDO

$$F_{\mu\nu} = \beta_{\mu\nu} + \sum_{[\sigma,\lambda]\in A} \left\{P_{\sigma\lambda}(\mu\nu|\sigma\lambda) - P^{\alpha}_{\sigma\lambda}(\mu\sigma|\nu\lambda)\right\}$$

For H and first row atoms,

$$F^{\alpha}_{\mu\nu} = \beta_{\mu\nu} + 2P_{\mu\nu}(\mu\nu|\mu\nu) - P^{\alpha}_{\mu\nu}[(\mu\mu|\nu\nu) + (\mu\nu|\mu\nu)]$$

CNDO

$$F^{\alpha}_{\mu\nu} = \beta_{\mu\nu} - P^{\alpha}_{\mu\nu}(\bar{\mu}\bar{\mu}|\bar{\nu}\bar{\nu})$$

Off-Diagonal: $\mu \in A$, $\nu \in B$, $A \neq B$ (most elements!)

NDDO

$$F^{\alpha}_{\mu\nu} = \beta_{\mu\nu} - \sum_{\substack{\sigma \in A \\ \lambda \in B}} P^{\alpha}_{\sigma\lambda} (\mu\sigma|\nu\lambda)$$

INDO

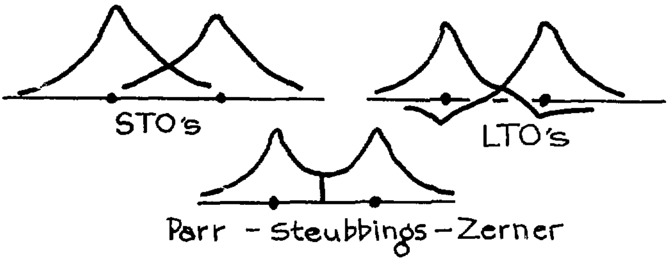
$$F^{\alpha}_{\mu\nu} = \beta_{\mu\nu} - P^{\alpha}_{\mu\nu}(\bar{\mu}\bar{\mu}|\bar{\nu}\bar{\nu})$$

CNDO

$$F^{\alpha}_{\mu\nu} = \beta_{\mu\nu} - P^{\alpha}_{\mu\nu}(\bar{\mu}\bar{\mu}|\bar{\nu}\bar{\nu})$$

Example: Integrals for N_2

	STO	LTO
$(2S_A 2S_A \| 2S_A 2S_A)$	0.709	0.738
$(2S_A 2S_A \| 2S_B 2S_B)$	0.452	0.437
$(2S_A 2S_A \| 2S_A 2S_B)$	0.277	-0.016
$(2S_A 2P_A \| 2S_A 2P_A)$	0.157	0.138
$(2S_A 2P^{\sigma}_A \| 2S^{\sigma}_B 2P^{\sigma}_B)$	0.121	0.084
$(2S_A 2P^{\sigma}_B \| 2P^{\sigma}_B 2P^{\sigma}_B)$	0.135	0.094



SOME FORMULATIONS

Resonance or Bonding Integrals

$$\beta_{\nu\mu}^{AA} = 0$$

$$\beta_{\nu\mu}^{AB} = (\beta_{\mu}^{A} + \beta_{\nu}^{B})\ \tilde{S}_{\mu\nu}/2$$

most often

$$\beta_{S}^{A} = \beta_{p}^{A} = \beta^{A}$$

$$\tilde{S}_{\mu\nu} = (\mu/\nu) = \Delta_{\mu\nu}$$

Nuclear Electronic Attraction

$$(\bar{\mu}|R_B^{-1}|\bar{\mu}) + V'_{\mu\mu}(c) + V_{\mu\mu}(STO \rightarrow LTO) \approx \gamma_{AB}$$

Core Integral

$$\left.\begin{matrix}\text{CNDO/1}\\ \text{INDO/1}\end{matrix}\right\}\quad \text{from ionization potential}$$

$$U_{\mu}^{CNDO/1} = -I_{\mu} - (Z^{V}-1)\gamma_{AA}$$

$$\left.\begin{matrix}\text{CNDO/2}\\ \text{INDO/2}\end{matrix}\right\}\quad \text{from } I_{\mu} + A_{\mu}$$

$$U_{\mu}^{CNDO/2} = -(I_{\mu} + A_{\mu})/2 - (Z^{V} - \tfrac{1}{2})\gamma_{AA}$$

<u>CNDO/1</u>

$$F^{\alpha}_{\mu\mu} = \underbrace{-I_{\mu}}_{\substack{\text{ionization}\\\text{potential of}\\\chi_\mu\text{ in free}\\\text{atom}}} + \underbrace{(P_{AA} - Z_A)\gamma_{AA}}_{\substack{\text{correc-}\\\text{tion for}\\\text{charged}\\\text{atom}}} + \underbrace{(1 - P^{\alpha}_{\mu\mu})\gamma_{AA}}_{\substack{\text{correction}\\\text{if not a}\\\text{"full" }\alpha\\\text{electron!}}} + \underbrace{\sum (P_{BB} - Z_B)\gamma_{AB}}_{\substack{\text{all charged}\\\text{neighbors}\\\text{have an}\\\text{influence!}}}$$

$$F^{\alpha}_{\mu\nu} = \beta_{\mu\nu} - \tfrac{1}{2} P^{\alpha}_{\mu\nu}\gamma_{\mu\nu}$$

where

$$\beta_{\mu\nu} = \frac{\Delta_{\mu\nu}}{2}\,(\beta^{A}_{\mu} + \beta^{b}_{\nu})$$

<u>MINDO/3</u> (Bingham, Dewar, Lo [JACS <u>97</u>, 1285 (1975)])

$$\gamma_{MN} = \frac{1}{\sqrt{R^2_{MN} + a^2_{MN}}} \qquad \text{(Ohno-Klopman)}$$

$$a_{mn} = \tfrac{1}{2}\left(\frac{1}{\gamma_{mm}} + \frac{1}{\gamma_{NN}}\right)$$

$$\beta_{\nu\mu} = (I_{\mu} + I_{\nu})S_{\mu\nu}\,B_{AB}$$

$$B_{AB} = \text{a "pair" parameter}$$

$$V_{NN} = \sum_{A\epsilon B} (CR)_{AB}$$

$$(CR)_{AB} = Z_A^V Z_B^V [\gamma_{AB} + (R_{AB}^{-} - \gamma_{AB}) f_3(R_{AB})]$$

$$f_3(R_{AB}) = \alpha_{AB} e^{-R_{AB}} \qquad \text{if NH and OH}$$

$$f_3(R_{AB}) = e^{-\alpha_{AB} R_{AB}} \qquad \text{otherwise}$$

Example: MNDO and MINDO/3 Predictions

Class of Compounds, etc.	*MNDO*	*MINDO/3*
ΔH_f (all compounds) Kcal/m	6.3	11.0
ΔH_f (HC) Kcal/m	6.0	9.7
ΔH_f (nitrogen compounds)	6.5	17.3
ΔH_f (oxygen compounds)	5.2	6.8
Bond lengths (all) Å	0.014	0.022
Bond lengths (CH) Å	0.009	0.019
Bond lengths (CC) Å	0.012	0.016
Bond lengths (NN) Å	0.032	0.074
Bond lengths (OO) Å	0.117	0.043
Bond angles (all)	2.8°	5.6°
about C	2.0°	4.4°
N	3.2°	7.1°
O	8.5°	10.7°

EXTENDED HUCKEL THEORIES

Consider Mulliken approximation

$$(\mu\nu|\sigma\sigma) = \frac{\Delta_{\mu\nu}}{2}\left\{(\mu\mu|\sigma\sigma) + (\nu\nu|\sigma\sigma)\right\}$$

$$(\mu\nu|R_B^{-1}) = \frac{\Delta_{\mu\nu}}{2}\left\{(\mu\mu|R_B^{-1}) + (\nu\nu|R_B^{-1})\right\}$$

<u>Diagonal Terms</u>

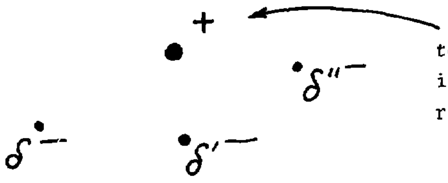
$$F_{\mu\mu} = U_{\mu\mu} - \sum_{B\neq A} Z_B(\mu\mu|R_B^{-1}) + \sum P_{\sigma\lambda}\left[(\mu\mu|\sigma\lambda) - \tfrac{1}{2}(\mu\sigma|\mu\lambda)\right]$$

$$= -I_\mu + (M_A - Z_A)\gamma_{AA} + (1 - \tfrac{1}{2}M_{\mu\mu})\gamma_{AA} + \sum_{B\neq A}(M_B - Z_B)\gamma_{\bar{A}\bar{B}}$$

simple Huckel or extended Huckel

iterative extended Huckel + higher order terms

the big error



too attractive if one ignores neighboring repulsions

Off-Diagonal Terms

$$H_{\mu\nu} = (H_{\mu\mu} + H_{\nu\nu})\Delta_{\mu\nu} K_{\mu\nu}/2$$

Usually, $K_{\mu\nu} = K = 1.7$ to 2.0, but different $K_{\mu\nu}$'s lead to improved results *if* one is careful with symmetry (the NEMO method of Newton and Libscomb). The method is then rotationally variant, *or*

$$H_{\mu\nu} = \sqrt{H_{\mu\mu}H_{\nu\nu}}\, \Delta_{\mu\nu}K$$

$$f = -\tfrac{1}{2}\nabla^2 + \sum_B V_B$$

$$f^A\chi_\mu = \varepsilon_\mu^A \chi_\mu$$

$$\begin{aligned}\varepsilon_\mu^A &= (\mu|-\nabla^2/2 + V_A|\mu) \qquad , \qquad \mu \in A \\ &= (\mu|-\nabla^2/2 - Z_A/R_A|\mu) + \sum_{\sigma\in A} M_{\sigma\sigma}(\overline{\sigma\mu}) \\ &= -I_\mu + R_\mu\end{aligned}$$

$$F_{\mu\mu} = \varepsilon_\mu^A + \sum_{B\neq A} (M_{BB} - Z_B^V)\gamma_{\bar{A}\bar{B}}$$

$$
\begin{aligned}
F_{\mu\nu} &= (\mu|f + \nabla^2/2 - \nabla^2/2|\mu) \\
&= (\mu|-\nabla^2/2 + V_A|\nu) + (\mu|-\nabla^2/2 + V_B|\nu) \\
&+ \tfrac{1}{2}(\mu|\nabla^2|\nu) + \sum_{C\neq A,B} (\mu|V_C|\nu) \\
&= (\epsilon_\mu^A + \epsilon_\mu^B)\Delta_{\mu\nu} + \tfrac{1}{2}(\mu|\nabla^2|\nu) \\
&+ \tfrac{1}{2}\,\Delta_{\mu\nu} \sum_{C\neq A,B} (M_{CC} - Z_C)(\gamma_{AC} + \gamma_{BC})
\end{aligned}
$$

Mulliken Approximation

$$\chi_\mu(1)\chi_\mu(1) = \frac{\Delta_{\mu\nu}}{2}\{\chi_\mu(1)\chi_\mu(1) + \chi_\nu(1)\chi_\nu(1)\}$$

$$\Rightarrow\ (\mu\nu|R_B^{-1}) = \frac{\Delta_{\mu\nu}}{2}[(\mu\mu|R_B^{-1}) + (\nu\nu|R_B^{-1})]$$

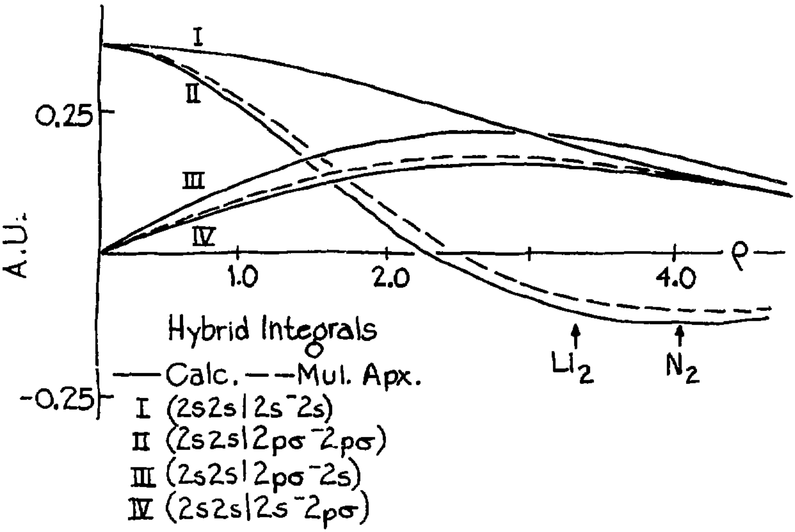
$$(\mu\nu|\sigma\sigma) = \frac{\Delta_{\mu\nu}}{2}[(\mu\mu|\sigma\sigma) + (\nu\nu|\sigma\sigma)]$$

Harris-Rein

$$\tilde{\Delta}_{\mu\nu} = \frac{4[(\mu\nu|R_B^{-1}) + (\mu\nu|R_A^{-1})]}{[(\mu\mu|R_A^{-1})+(\mu\mu|R_B^{-1})+(\nu\nu|R_A^{-1})+(\nu\nu|R_B^{-1})]}$$

$$\tilde{\tilde{\Delta}}_{\mu\nu} = \frac{4[(\mu\nu|\bar{\mu}\bar{\mu}) + (\mu\nu|\bar{\nu}\bar{\nu})]}{[(\mu\mu|\mu\mu)+(\nu\nu|\nu\nu)+(\mu\mu|\nu\nu)+(\nu\nu|\mu\mu)]}$$

$$\rightarrow\ J_{\mu\nu} = \tilde{\tilde{\Delta}}_{\mu\nu}(J_{\mu\mu} + J_{\nu\nu})$$


I
II
III
IV
0.25
-0.25
A.U.
1.0
2.0
4.0
ρ
Li2
N2
Hybrid Integrals
—Calc. - - -Mul. Apx.
I (2s2s|2s⁻2s)
II (2s2s|2pσ⁻2pσ)
III (2s2s|2pσ⁻2s)
IV (2s2s|2s⁻2pσ)

Example: Performance Examples

	Relative Error	Relative Time	Cost Efficiency
INDO	14	0.18	0.4
PRDDO	1	1	1
STO-3G	0.2	16	0.3
AAMOM	2.5	0.4	1
VRDDO	~0.2	~12	~0.4
VRDDO MODPOT	~0.5	~5	~0.4
ESE MO/2C	~7	~0.7	~0.2

$$+ \quad \frac{1}{(\text{Relative error}) \times (\text{Relative cost})}$$

(from Halgren and Lobscomb, et al., [JACS 100, 6595 (1978)]

MOLECULAR ORBITAL THEORY "REVISITED"

1) $\underline{H\Psi = E\Psi}$

H is non-relativistic time-independent fixed nuclei Hamiltonian

2) <u>MO Approximation</u>

$$\Psi(\vec{1}\ldots\vec{n}) \approx \psi_o(\vec{1},\ldots\vec{n}) \underset{\text{MO Apx}}{=} \mathcal{O}_s\mathcal{A}[\underbrace{\overset{\text{MO's}}{\phi_1}(1)\,\phi_2(2)\ldots\phi_n(n)}_{\text{electronic configuration}}]$$

"The Big Approximation"

3) <u>LCAO-MO Approximation</u>

$$\phi_i = \sum_{\mu=1}^{n} \chi_\mu\, C_{\mu i} = \boldsymbol{\chi}\, \mathbb{C}_i$$

4) <u>Variational Principle</u>

$$W \equiv \frac{\langle \psi_T | H | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} \geqslant E_x$$

$$\delta W = 0 \Rightarrow (\mathbb{F} - \varepsilon_i \boldsymbol{\Delta})\, \mathbb{C}_i^+ = 0$$

(secular equation)

$$\boldsymbol{\Delta} = \boldsymbol{\chi}^+\boldsymbol{\chi} = \text{overlap}$$

$$\mathbb{F} = \boldsymbol{\chi}^+ F \boldsymbol{\chi} = \langle \boldsymbol{\chi} | F | \boldsymbol{\chi} \rangle$$

$$F^{\alpha}_{\mu\nu} = \langle \chi_\mu | F | \chi_\nu \rangle = h_{\mu\nu} + J_{\mu\nu} - K^{\alpha}_{\mu\nu} \qquad \text{(UHF)}$$

$$= t_{\mu\nu} - \sum_A Z_A(\mu|R_A^{-1}|\nu) + \sum P_{\sigma\lambda}(\mu\nu|\sigma\lambda)$$

$$- \sum P^{\alpha}_{\sigma\lambda}(\mu\sigma|\nu\lambda)$$

where

$$t_{\mu\nu} = \text{kinetic energy} = -\tfrac{1}{2}(\mu|\nabla^2|\nu)$$

$$(\mu|R_B^{-1}|\nu) = (\mu\nu|R_B^{-1}) = \int d\tau(1)\ \chi^*_\mu(1)\chi_\nu(1)R_B^{-1}$$

$$(\mu\nu|\sigma\lambda) = \int d\tau(1)d\tau(2)\ \chi_\mu(1)\chi_\nu(1)\ \frac{1}{r_{12}}\ \chi_\sigma(2)\chi_\lambda(2)$$

$$P^{\alpha}_{\mu\nu} = \sum_a^{\text{MO's}} C^{\alpha}_{\mu a}\ C^{\alpha}_{\nu a} n^{\alpha}_a \qquad (n_a = \text{occupancy} = 0 \text{ or } 1)$$

= Fock Dirac "α" density

$$\underline{P} = \underline{P}^{\alpha} + P^{\beta}$$

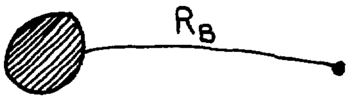
$$F^{\alpha}_{\mu\mu} = U_{\mu\mu} - \sum_{B\neq A} (\mu|Z_B/R_B|\mu) + \sum_{\sigma,\lambda} P_{\sigma\lambda}(\mu\mu|\sigma\lambda) - \sum P^{\alpha}_{\sigma\lambda}(\mu\sigma|\mu\lambda)$$

A) $U_{\mu\mu}$ = "core integral" = $(\mu|-\tfrac{1}{2}\nabla^2 - Z_A/R_A|\mu)$

an atomic-like integral

B) One electron two center integrals nuclear

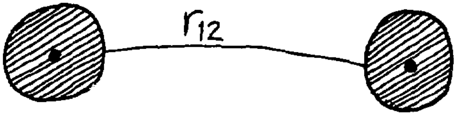
attraction $(\mu|Z_B/R_B|\mu)$



$\chi_\mu^*(1)\chi_\mu(1)$

C) Two-center two-electron

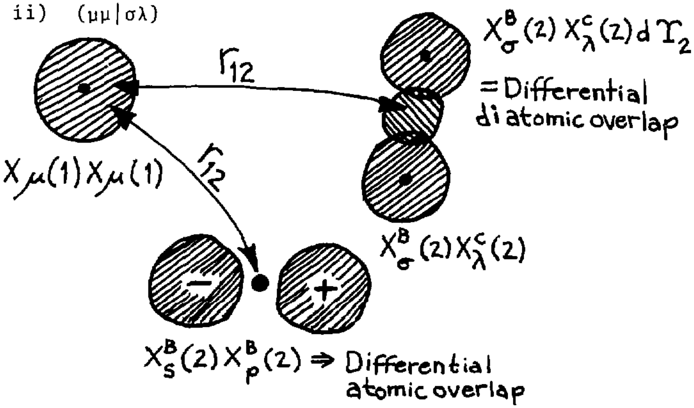
i) $(\mu\mu|\nu\nu)$



$\chi_\mu(1)\chi_\mu(1)$ $\chi_\nu(2)^*\chi_\nu(2)$

Coulomb interaction between two charge distributions with the "test" electron in $\chi_\mu^*(1)\chi_\mu(1)$ and $P_{\nu\nu}$ electrons in the other.

ii) $(\mu\mu|\sigma\lambda)$



Returning to $F^{\alpha}_{\mu\mu}$, we can write

$$F^{\alpha}_{\mu\mu} = U_{\mu\mu} - \sum_{B} (\mu|Z_B/R_B|\mu) + \sum_{a}^{MO} (\mu\mu|\phi_a\phi_a) - \sum_{a}^{MO} (\mu\phi^{\alpha}_a|\mu\phi_a)$$

so $\qquad \underline{\mathbb{F}}_{\mu}(\{\phi_a\}) \qquad$ or $\qquad \mathbb{F}(\mathbb{C})$

5) <u>Self-Consistent Field SCF</u>

Guess $\mathbb{C}_o$

$\mathbb{C}_o \rightarrow \mathbb{F}(\mathbb{C}_o) \rightarrow \mathbb{C}_1 \rightarrow \mathbb{C}_1 \rightarrow \mathbb{F}(\mathbb{C}_1) \rightarrow \mathbb{C}_2 \rightarrow$ etc.

CRITIQUE OF MOLECULAR ORBITAL THEORY

A. Computational difficulties

1) N^4 integrals $(\mu\nu|\sigma\lambda)$ where N is size of basis

2) N^3 matrix problem — solution of secular equation

All MO methods are limited "spiritually" by N^3 matrix problem.

Integral approximations try to reduce integral problem to N^3 or less

$$\begin{array}{ccccc} N^4 & \longrightarrow & N^3 & \longrightarrow & N^2 \\ & & \uparrow & & \\ & & \text{AAMOM} & & \text{differential overlap methods} \\ & & \text{PRDDO} & & \text{CNDO, INDO, NDDO, PPP} \\ & & & & \text{extended Huckel methods} \end{array}$$

B. Theoretical Limitations

Fundamentally incapable of yielding exact answers except for one electron case!

RETURN TO ATOMIC HARTREE-FOCK THEORY

1) Good intuitive feel for atoms as well as Hartree-Fock procedure.

2) Most approximate methods have parameters derived from atomic information.

Consider, for example

$$\psi = {}^2P = |s\bar{s}p|$$

$$E({}^2P) = 2U_s + U_p + J_{ss} + 2J_{sp} - K_{sp}$$

$$J_{sp} = (ss/pp) \quad , \quad J_{ss} = (ss/ss)$$

$$K_{sp} = (sp/sp)$$

Consider also a basis set of STO's (for now!!)

$$n\ell m = \chi_{n\ell m} = R_{n\ell}(r)\, Y_\ell^m(\theta\phi)$$

$$R_{n\ell}(r) = \eta_{m\ell}\, e^{-\xi r}\, r^{n-1} \quad , \text{ STO}$$

$$Y_\ell^m(\theta\phi) = P_\ell^m(\cos\theta)\, e^{im\phi}$$

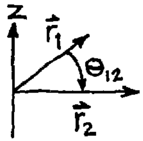
with such a basis one-center integrals are easy!

$$(1s|t|1s) = \xi_1/2 \quad , \quad (2s|t|2s) = \xi_{2s}/6$$

$$(2p|t|2p) = \xi_{2p}/2 \quad , \quad (\chi_{n\ell m}|Z_A/R_A|\chi_{n\ell m}) = \frac{Z_A\xi_{n\ell}}{n}$$

Two-center integrals are not quite as easy, but

$$\frac{1}{r_{12}} = \sum_{K=0}^{\infty} \frac{r_<^{k}}{r_>^{k+1}} P_k(\cos\theta_{12})$$



$$\frac{1}{r_{12}} = \sum_{k=0}^{\infty} \sum_{m=-k}^{k} \frac{(k-|m|)!}{(k+|m|)!} \frac{r_<^{k}}{r_>^{k+1}} P_k^m(\theta_1)P_k^m(\theta_2)\, e^{ik(\phi_1-\phi_2)}$$

$$(ij|k\ell) = \sum_{\mu}^{\infty} R^{\mu}(ijk\ell)\int d\Omega_1 d\Omega_2'\; Y_i^*(1)Y_j(1)Y_k^*(2)Y_\ell(2)$$

$$\times \quad P_\mu^m(1)P_\mu^m(2)\; e^{im(\phi_1-\phi_2)}$$

$$R^{\mu}(ij|k\ell) = \begin{matrix}\text{Slater Condon}\\ \text{factor}\end{matrix} = \int r_1^2 dr_1 \int r_2^2 dr_1 \;\; R_i^*(r_1)R_j(r_2)$$

$$\times \quad \frac{r_<^{\mu}}{r_>^{\mu+1}} R_k(r_2)^* R_\ell(r_2)$$

$$(ij|k\ell) = \delta_{s_i s_j}\, \delta_{s_k s_\ell}\, \delta(m_i+m_k, m_j+m_\ell)$$

$$\sum_{k=0} \underbrace{C^k(\ell_i m_i \ell_j m_j)C^k(\ell_k m_k \ell_\ell m_\ell)}_{\text{Clebsch-Gordon coefficients}}$$

The sum is not infinite, but to $k = \inf(\ell_i+\ell_j, \ell_k+\ell_\ell)$

special cases!

$$(ii|jj) = \sum_k a^k(\ell_i m_i \ell_j \ell_j)\ F^k(ij)$$

where

$$F^k(ij) = R^k(ii|jj) \quad ,$$

$$a^k(\ell_i m_i \ell_j m_j) = C^k(\ell_i m_i \ell_i m_i) C^k(\ell_j m_j \ell_j m_j)$$

$$(ij|ij) = \sum_k b^k(\ell_i m_i \ell_j m_j)\ G^k(ij)$$

$$G^k(ij) = R^k(ijij) \quad ,$$

$$b^k(\ell_i m_i \ell_j m_j) = C^k(\ell_i m_i \ell_j m_j)^2 \quad .$$

Now for only s and p orbitals

$$J_{ss} = F^o(ss) \quad , \qquad J_{sp} = F^o(sp) \quad ,$$

$$K_{sp} = G'(sp)/3 \quad , \qquad J_{xx} = F^o(pp) + \frac{2}{25} F^2(pp) \quad ,$$

$$J_{xy} = F^o(pp) - \frac{4}{25} F^2 pp \quad , \qquad K_{xy} = \frac{3}{25} F^2(pp) \quad .$$

For our example of $^2P = |s\bar{s}p|$

$$E(^2P) = 2U_s + U_p + F^o(ss) + 2F^o(sp) - 1/3\ G'(sp)$$

This is a great theory! Can derive energy for any atomic spectroscopic state providing

$$S^2\psi = S(S+1)\psi$$

$$L^2\psi = L(L+1)\psi$$

$$S_z\psi = m_s\psi$$

$$L_z\psi = m_\ell\psi$$

For now consider average energies of a pair of electrons (two electron part).

$$(\overline{ss}) = (s\bar{s}) = F^0(ss)$$

$$(\overline{sp}) = \tfrac{1}{4}\{(sp)+(\bar{s}\bar{p})+(s\bar{p})+(\bar{s}p)\} = \tfrac{1}{2}\{(sp)+(s\bar{p})\}$$

$$= F_0(sp) - \tfrac{1}{6}G'(sp)$$

$$(\overline{pp}) = \frac{1}{15}\ (15\ \text{possible}) = F^0(pp) - \frac{2}{25}F^2(pp)$$

For our case

$$E(^2P) = 2U_s + U_p + (\overline{ss}) + 2(\overline{sp})$$

and, in general, the average energy of a configuration is given by

$$E_{ave}(s^\ell p^m d^n) = \ell U_s + mU_p + nU_d + \sum(\overline{pair}) .$$

Consider, as a typical atomic property, the ionization energy I_μ

$$I_\mu = E(+ve) - E(atom)$$

$$I_s = E(s^{\ell-1}p^m d^n) - E(s^\ell p^m d^n)$$

$$= U_{ss} - (\ell-1)F^o(ss) - m\left[F^o(sp) - \frac{1}{6}G'(sp)\right]$$

$$- \frac{n}{2}\left[F^o(sd) - \frac{G^2(sd)}{10}\right]$$

$$I_p = -U_{pp} - (m-1)\left[F^o(pp) - \frac{2}{25}F^2(pp)\right]$$

$$- \ell\left[F^o(sp) - \frac{G^1(sp)}{6}\right] - m\left[F^o(pd) - \frac{G^1(pd)}{15} - \frac{3G^3(pd)}{70}\right]$$

$$I_d = -U_{dd} - (n-1)\left[F^o(dd) - \frac{2}{63}(F^2(dd)+F^4(dd))\right]$$

$$- \ell\left[F^o(sd) - \frac{G^2(sd)}{10}\right] - m\left[F_o(pd) - \frac{G^1(pd)}{15} - \frac{3}{70}G^3(pd)\right]$$

$$A_\mu = E(atom) - E(ve) \quad , \qquad etc.!$$

One can calculate from these expressions ionization energies, but it is far more common to estimate the core integral from atomic information, and then use this $U_{\mu\mu}$ in molecular calculations, i.e. ,

$I_i \Rightarrow U_i$	CNDO/1	Hückel
	INDO/1	EHT
	PPP	most others

$(I_i+A_i)/2 \Rightarrow U_i$ CNDO/2
INDO/2

can also obtain this information from atomic spectroscopy in a similar fashion.

Several "fakes":

1) Minimum basis set representation for atom.
2) Frozen orbital representation for positive and negative ions.
3) What happened to the inner shell orbitals?

CORE VALENCE SEPARATION

Do we need the core electrons?

a) Chemists seldom consider inner-shell electrons for *most* chemical phenomena
b) Early calculations in which valence orbitals were *orthogonalized* to the core showed that in some sense core orbitals were separable.

Why would we want to do this?

Minimum basis set for benzene C_6H_6: 36 a.o.'s but 30 valence

$$\left(\frac{30}{36}\right)^4 \sim \text{ half the number of integrals}$$

Double-ξ for $CuCl_2$ (78 functions) of which 34 are valence.

$${\frac{34}{78}}^{4} \sim \frac{1}{16} \text{ the number of integrals!}$$

Cannot just drop core orbitals by "wishful thinking" however.

For N_2,

$$F_\sigma = \begin{array}{cc|ccc|} & & 1s & 2s & 2p\sigma \\ \hline A & 1s & -15.71 & & \\ & 2s & -3.78 & -2.04 & \\ & 2p\sigma & -0.03 & -0.35 & -0.82 \\ B & 1s & -0.00 & -0.88 & -1.48 \\ & 2s & -0.88 & -1.17 & -1.05 \\ & 2p\sigma & -1.98 & -1.05 & -0.53 \end{array}$$

$F_{1s_A 2s_A} = -3.78$ is second largest number in F matrix.

Cannot hope to just toss it away!

EVERYONE'S CORE "POTENTIAL"

Use partitioning technique:

$$(\mathbf{F} - \mathbf{E}\boldsymbol{\Delta})\mathbf{C} \qquad \mathbf{MC} = 0$$

$$\begin{vmatrix} \mathbf{M}_{cc} & \mathbf{M}_{cv} \\ \mathbf{M}_{vc} & \mathbf{M}_{vv} \end{vmatrix} \begin{vmatrix} \mathbf{C}_c \\ \mathbf{C}_v \end{vmatrix} = 0$$

$$\mathbf{M}_{cc}\mathbf{C}_c + \mathbf{M}_{cv}\mathbf{C}_v = 0 \Longrightarrow \mathbf{C}_c = -\mathbf{M}_{cc}^{-1}\,\mathbf{M}_{cv}\mathbf{C}_v$$

$$\mathbf{M}_{vc}\mathbf{C}_c + \mathbf{M}_{vv}\mathbf{C}_v = 0$$

$$\left(\mathbf{M}_{vv} - \underbrace{\mathbf{M}_{vc}\mathbf{M}_{cc}^{-1}\,\mathbf{M}_{cv}}_{-V_{cc}}\right)\mathbf{C}_v = 0$$

This yields:

$$\left(\mathbb{H}'_{vv} - E\mathbb{\Delta}_{vv}\right)\mathbf{C}_v = 0$$

$$\mathbb{H}'_{vv} = \mathbb{H}_{vv} + \mathbf{V}_{vv}$$

where $\mathbf{V}_{vv}$ is an exact core potential.

$$-V_{ij} = \sum_{\alpha,\beta} (F_{i\alpha} - \Delta_{i\alpha}E)(\mathbb{F}_{cc} - \mathbb{\Delta}_{cc}E)^{-1}_{\alpha\beta}(F_{\beta j} - \Delta_{\beta j}E)$$

where $i,j,\ldots$ are valence a.o.'s and $\alpha,\beta,\ldots$ are core a.o.'s. The problem is now more difficult than when we started! *But* we note that M_{cc} is nearly diagonal (see, for example $\mathbb{F}$ for N_2: the $F_{1s_A 1s_B} \approx 0.00$), so

$$\mathbf{M}_{cc} = \mathbb{A} + \mathbb{B}$$

where $\mathbb{A}$ is diagonal and $\mathbb{B}$ is off-diagonal

$$M_{cc}^{-1} = (\mathbb{A}+\mathbb{B})^{-1} = \mathbb{A}^{-1} - \mathbb{A}^{-1}\,\mathbb{B}(\mathbb{A}+\mathbb{B})^{-1}$$

{We can check this by multiplying on the right by (A+B) and deriving 1=1}, and iterate

$$(\mathbb{A}+\mathbb{B})^{-1} = \mathbb{A}^{-1} - \mathbb{A}^{-1}\mathbb{B}\mathbb{A} + \mathbb{A}^{-1}\mathbb{B}\mathbb{A}^{-1}\mathbb{B}\mathbb{A}^{-1} - \dots$$

so

$$-V_{ij} = \sum_{\alpha,\beta} (F_{i\alpha} - \Delta_{i\alpha}E)\underbrace{(\mathbb{M}_{cc})^{-1}_{\alpha\beta}}_{\text{expand}}(F_{\alpha j} - \Delta_j E) = \sum_{k=1} V_{ij}^{(k)}$$

where the above equation is a Brillouin Wigner perturbation sequence

$$-V_{ij}^{(1)} = \left(\mathbb{M}_{vc}\, \mathbb{A}_{cc}^{-1}\, \mathbb{M}_{cv}\right)_{ij}$$

$$-V_{ij}^{(2)} = \left(\mathbb{M}_{vc}\, \mathbb{A}_{cc}^{-1}\, \mathbb{B}_{cc}\mathbb{A}_{cc}^{-1}\, \mathbb{M}_{cv}\right)_{ij}$$

etc.

Consider

$$F_{i\alpha} = \Delta_{i\alpha}F_{\alpha\alpha} + G_{i\alpha}$$

As an empirical observation, $G_{i\alpha}$ is small! Then,

$$-V_{ij}^{(1)} = (\mathbb{M}_{vc}\, \mathbb{A}_{cc}^{-1}\, \mathbb{M}_{cv})_{ij}$$

$$= \sum (F_{i\alpha} - \Delta_{i\alpha}E)(F_{\alpha\alpha}-E)^{-1}(F_{\alpha j} - E\Delta_{\alpha j})$$

$$= \left\{\sum_{\alpha} \left[\Delta_{i\alpha}(F_{\alpha\alpha}-E)\Delta_{\alpha j} + G_{i\alpha}\Delta_{\alpha j} + \Delta_{i\alpha}G_{\alpha j} + \frac{G_{i\alpha}G_{\alpha j}}{(F_{\alpha\alpha}-E)}\right]\right\}$$

$$-V_{ij}^{(2)} = \text{etc.}$$

What if $G_{i\alpha} = 0$ *all* i? Then, $\underline{F}\chi_\alpha = F_{\alpha\alpha}\chi_\alpha$ or each χ_α is an eigenfunction of the Fock operator $\underline{F}$. Empirically this is very nearly so! If $G_{i\alpha} = 0$, then M_{cc} *is* diagonal *and*

$$-V_{ij} = -V_{ij}^{(1)} = \sum_\alpha \Delta_{i\alpha}(F_{\alpha\alpha} - E)\Delta_{\alpha j}$$

$$= \text{Phillips Kleinman "pseudo-potential"}$$

This is reasonably accurate, *but* depends on E, the valence orbital eigenvalues, and

a) Must solve iteratively for each $E_{valence}$.

b) Each orbital solves a different $\underline{F}$ operator.

c) Too many disadvantages!

so, note that $\varepsilon_\alpha \approx F_{\alpha\alpha} < E_{valence}$ (i.e., for N_2, $\varepsilon_\alpha \approx -15.7$ and $\varepsilon_{valence} \sim -1.0$). Then,

$$-V_{ij} \approx \sum_\alpha \Delta_{i\alpha}\, \tilde{\varepsilon}_\alpha \Delta_{\alpha j}$$

which is related to everyone's "pseudo-potential", where $\tilde{\varepsilon}_\alpha$ is a parameter and $\Delta_{i\alpha}$ is an inner-shell outer-shell overlap often estimated by an effective potential.

For now, note that

$$-V_{ii} = \underbrace{\Delta_{i\alpha}^2 \tilde{\varepsilon}_\alpha}_{\text{one center}} + \sum_{\alpha \neq \beta} \underbrace{\Delta_{i\beta}^2 \tilde{\varepsilon}_\beta}_{\text{two center}} \qquad i,\alpha \in A$$

$$U'_{ii} = U_{ii} + \Delta^2_{i\alpha}\,\tilde{\varepsilon}_\alpha$$

This last expression is what came from experimental information when we ignored the core. Thus, empirical methods that utilize atomic information for core integrals implicitly include the inner shell. The two-center part of this repulsion must then be included parametrically – usually by scaling two-center nuclear attraction integrals!

PI-SIGMA (PEEL-CORE) SEPARATION

Consider planar molecular and two-elements of symmetry: E (does nothing) and σ_h (reflects in plane). So,

$$[E,H] = [\sigma_h,H] = 0$$

because E and σ_h cannot change any observable property of the system, especially the energy! Usually $[h_{eff},g_i] = 0$ whenever $[H,g_i] = 0$ by Roothaan-Hall construction. Then,

$$h_{eff}(1)\phi_i(1) = \varepsilon_i\ \phi_i(1)$$

$$E\phi_i = \phi_i$$

$$\sigma_h\phi_i = \lambda\phi_i$$

(since $[\sigma_h, h_{eff}] = 0$)

$$\sigma_h^2\phi_i = \lambda\sigma_h\phi_i = \lambda^2\phi$$

$$\sigma_h\phi_i = E\phi_i = \phi_i$$

The $\sigma_h\phi_i$ *equals* $E\phi_i$ since reflecting and reflecting back does nothing.

$$\Rightarrow \quad \lambda^2 = 1$$

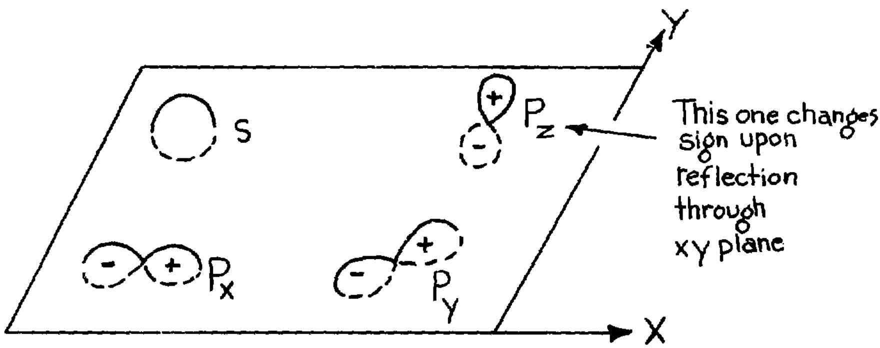
$$\lambda = \pm 1$$

$$\sigma_h\phi_i \underset{\pi\text{ MO's}}{\overset{\sigma\text{ MO's}}{=}} \pm\phi_i$$

Now, LCAO-MO

$$\phi_i = \sum_\mu \underbrace{\chi_\mu}_{\text{ao's}} C_{\mu i}$$

$$\sigma_h\phi_i = \pm\phi_i = \sum_\mu (\sigma_h\chi_\mu)C_{\mu i}$$



$$\sigma_h(s, p_x, p_y, d_{x^2-y^2}, d_{xy}, d_z \;:\; p_z, d_{xz}, d_{yz})$$

$$= \underbrace{(s, p_x, p_y, d_{x^2-y^2}, d_{xy}, d_z}_{\sigma \text{ ao's}} \;:\; \underbrace{-p_z, -d_{xz}, -d_{yz})}_{\substack{\pi \text{ ao's have} \\ \text{changed sign}}}$$

$$\sigma_n \phi_i = \pm\phi_i = \pm \sum_{\mu}^{(\sigma)} \chi_\mu^\sigma C_{\mu i} \pm \sum_{\nu}^{(\pi)} \chi_\nu^\pi C_{\nu i}$$

$$= \sum (\sigma_n \chi_\mu) C_{\mu i} = \underset{\sigma \text{ ao's}}{\pm} \sum_{\mu}^{(\sigma)} \chi_\mu^\sigma C_{\mu i} \underset{\pi \text{ ao's}}{-} \sum_{\nu}^{\pi} \chi_\nu^\pi C_{\nu i}$$

Comparing coefficients implies:

π MO's have only π symmetry ao's

σ MO's have only σ symmetry ao's

$$(\mathbb{F} - \varepsilon_i \Delta)\mathbb{C}_i = \mathbb{M}\mathbb{C}_i = \begin{vmatrix} \mathbb{M}_{\sigma\sigma} & 0 \\ 0 & \mathbb{M}_{\pi\pi} \end{vmatrix} \begin{vmatrix} \mathbb{C}_{\sigma i} \\ \mathbb{C}_{\pi i} \end{vmatrix}$$

These elements must be zero or σ MO's would have π ao's and π MO's would have σ ao's.

$\{$ σ ao's transform as a' irreducible representation of C_s $\}$

$\{$ π ao's as a", H as a', and $\langle \mu | H | \nu \rangle$ must transform as a', etc. $\}$

Note that this block diagonal form does not imply that $\mathbf{M}_{\pi\pi}$ does not depend on σ electrons — it does!

Following Lykos and Parr:

$$\psi = \mathcal{A}(\Sigma)(\pi) = |(\Sigma)(\pi)|$$

$$\langle\psi|\psi\rangle = (\Sigma)|(\Sigma)\rangle = \langle(\pi)|(\pi)\rangle = 1$$

We want

$$\langle|(\Sigma)(\pi)| \ |H| \ |(\Sigma)(\pi)|\rangle = E_{e\ell} = E_\sigma + E_\pi$$

with

$$E_\sigma = \langle(\Sigma)|H_\sigma|(\Sigma)\rangle$$

$$E_\pi = \langle(\pi)|H_\pi|(\pi)\rangle$$

Can these conditions be met? Of course!

One solution

$$H_\sigma = \sum_{\sigma=1}^{n_\sigma} h(\sigma) + \tfrac{1}{2}\sum_{\sigma,\lambda}^{n_\sigma} \frac{1}{r_{\sigma\lambda}}$$

$$H_\pi = \sum_{\mu=n_\sigma+1}^{n_\sigma+n_\pi} h_{core}(\mu) + \tfrac{1}{2}\sum_{\mu,\nu}^{n_\sigma+n_\pi} \frac{1}{r_{\mu\nu}}$$

where

$$h(\sigma) = (\sigma|-\tfrac{1}{2}\nabla^2|\sigma) - \sum_A (\sigma|Z_A/R_A|\sigma)$$

$$h_{core}(\mu) = h(\mu) + J_\sigma(\mu) - K_\sigma(\mu)$$

$$J_\sigma(\mu)(\pi) = \int d\tau_\sigma(\Sigma) \sum_\sigma \frac{1}{r_{\sigma\mu}} (\Sigma)(\pi)$$

represents Coulomb repulsion between σ and π electrons, and

$$K_\sigma(\mu)(\pi) = \int d\tau_\sigma(\Sigma) \sum_\sigma \frac{P_{\sigma\pi}}{r_{\sigma\mu}} (\Sigma)(\pi)$$

represents the exchange term between sigma and pi electrons. The above separation is *exact* and one could iterate, first solving for (Σ), then (π), etc.

Approximations

1) $$(\pi'|K_\sigma|\pi) = \sum_\sigma (\pi'\sigma|\pi\sigma)$$

$$= \sum_\sigma \sum_{\substack{\mu\nu \\ \lambda,\delta}} C_{\mu\pi'} C_{\nu\sigma} C_{\lambda\pi} C_{\delta\sigma} (\mu\nu|\lambda\delta) \approx 0$$

Larger terms are when $\mu=\nu$ $(\mu\mu|\lambda\lambda)$, but $C_{\mu\pi'}C_{\mu\sigma} = 0$ for an a.o. cannot be both a component of a σ and a π MO!

2) $$(\pi|J_\sigma|\pi) = \sum_\sigma (\pi\pi|\sigma\sigma) = \sum_\sigma \sum_{\mu\nu} C_{\mu\sigma} C_{\nu\sigma} (\pi\pi|\mu\nu)$$

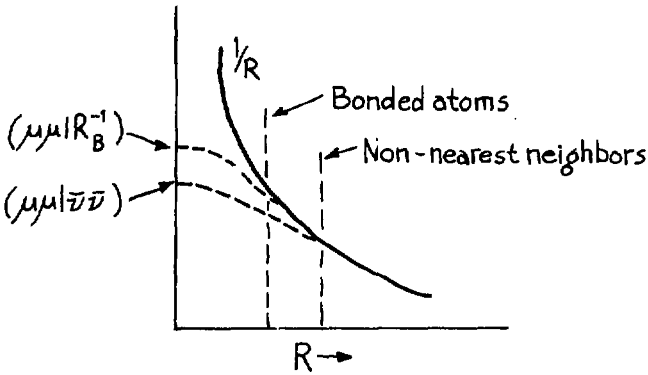
$$\approx \sum_\mu M_{\mu\mu} (\pi\pi|\mu\mu) = \sum_A M_A^\sigma (\pi\pi|\bar{\mu}\bar{\mu})$$

Mulliken integral approximation   Mulliken orbital population

M_A^σ represents the number of σ electrons from

$$(\pi|h|\pi) = (\pi|t|\pi) - \sum_A (\pi|Z_A/R_A|\pi)$$

$$(\pi|h_{core}|\pi) = (\pi|t|\pi) + \sum_A M_A^\sigma(\pi\pi|\bar{\mu}\bar{\mu}) - Z_A(\pi|R_A^{-1}|\pi)$$



$$(\pi\pi|\bar{\mu}\bar{\mu}) \approx (\pi|R_B^{-1}|\pi)$$

$$\Rightarrow \quad h_{core} = -\tfrac{1}{2}\nabla^2 - \sum_A \underbrace{(Z_A - n_A^\sigma)}_{\equiv Z_A^\pi} R_A^{-1}$$

where n_A^σ is the number of σ electrons of atom A.

Pariser Parr Pople (PPP) Theory

$$\phi_i = \sum_\mu \chi_\mu^\pi C_{\mu i}$$

$$(\mathbf{F} - \boldsymbol{\epsilon}_i \boldsymbol{\Delta})\mathbf{C}_i = 0$$

$$\boldsymbol{\Delta} = \mathbf{1} \leftrightarrow (\chi_\mu^\pi | \chi_\nu^\pi) = \delta_{\mu\nu}$$

$$\mathbf{FC} = \mathbf{C}\boldsymbol{\epsilon} \rightarrow \mathbf{C}^+\mathbf{FC} = \boldsymbol{\epsilon}$$

$$F_{\mu\mu} = U_{\mu\mu} + \sum_\sigma (P_{\sigma\sigma} - Z_\sigma^\pi)\gamma_{\mu\sigma} - \tfrac{1}{2}P_{\mu\mu}\gamma_{\mu\mu}$$

$$\gamma_{\mu\nu} \equiv (\mu\mu|\nu\nu)$$

$$U_{\mu\mu} = \alpha_\mu = -I_\mu \qquad \text{or} \qquad -\frac{(I_\mu + A_\mu)}{2} - \frac{\gamma_{\mu\mu}}{2}$$

$$F_{\mu\nu} = \beta_{\mu\nu} - \tfrac{1}{2}P_{\mu\nu}\gamma_{\mu\nu}$$

$\beta_{\mu\nu}$ is a parameter usually chosen to fit spectra after a singles only CI. More about PPP later.

AB INITIO HARTREE FOCK

Lecture 6/7

by

John A. Pople
Department of Chemistry
Carnegie-Mellon University
Pittsburgh, Pennsylvania

GAUSSIAN PROGRAMS

G 70 (IBM)

S,P bases + RHF + UHF

G 76 (CDC)

S,P,D bases + RHF + UHF

G 78 (DEC-VAX)

G 76 + direct minimization SCF

+ energy derivatives (OPT)

+ correlation by MP2, MP3

G 80 (DEC-VAX)

G 78 + CID + CISD + archive

THEORETICAL MODEL CHEMISTRY

REQUIREMENTS

1. Uniqueness and universality
2. Simplicity
3. Interpretability
4. Size consistency (i.e., additivity for isolated systems).

HARTREE-FOCK THEORY

$$\Psi = (n!)^{-\frac{1}{2}} |\chi_1 \chi_2 \ldots \chi_n|$$

Spinorbitals $\chi_i = \sum^{N} c_{\mu i} \, \omega_\mu$

Basis Functions ω_μ

Coefficients $c_{\mu i}$ adjusted to minimize

$$E = \int \Psi^* H \Psi \, d\tau$$

RHF THEORY

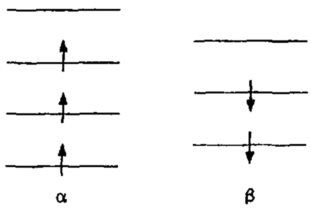
closed shell open shell

Spinorbitals $\chi = \psi \times \begin{cases}\alpha \\ \beta\end{cases}$

Only one set of coefficients $c_{\mu i}$.

Advantage: eigenfunction of $\underset{\sim}{S}^2$

UHF THEORY



Spinorbitals $X = \psi^{\alpha}\alpha$ or $\psi^{\beta}\beta$

Two sets of coefficients $c^{\alpha}_{\mu i}$, $c^{\beta}_{\mu i}$

Advantages:

more flexible

size-consistent (dissociation)

Disadvantages:

not an eigenfunction of S^2

ROOTHAAN EQUATIONS

$$\sum_{\nu} (F_{\mu\nu} - \varepsilon_i \, \delta_{\mu\nu}) c_{\nu i} = 0$$

$$F_{\mu\nu} = H^{\text{core}}_{\mu} + \sum_{\lambda\sigma} P_{\lambda\sigma} [(\mu\nu|\lambda\sigma) - \tfrac{1}{2}(\mu\lambda|\nu\sigma)]$$

$$P_{\lambda\sigma} = \text{density matrix}$$

$$= 2 \sum_{i}^{n/2} c^{*}_{\lambda i} \, c_{\sigma i}$$

Overlap : $S_{\mu\nu} = \int \phi_\mu \phi_\nu \, d\tau$

Core H : $H_{\mu\nu}^{core} = \int \phi_\mu \, H^{core} \, \phi_\nu \, d\tau$

$$(\mu\nu|\lambda\sigma) = \iint \phi_\mu(1)\,\phi_\nu(1)\,\frac{1}{r_{12}}\,\phi_\lambda(2)\,\phi_\sigma(2)\;d\tau_1 d\tau_2$$

<u>MATHEMATICAL FORM OF BASIS FUNCTIONS</u>

Slater : $e^{-\xi r}$, $xe^{-\xi r}$, etc.

Advantages: Like AO's

Disadvantages: $(\mu\nu|\lambda\sigma)$ hard

Gaussian (Boys) : $e^{-\alpha r^2}$, $xe^{-\alpha r^2}$, etc.

Advantages: $(\mu\nu|\lambda\sigma)$ easy

Disadvantages: Contraction usually necessary

<u>EVALUATION OF INTEGRALS</u>

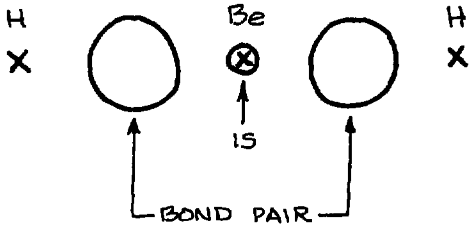
Boys (1950): uses $\int_0^1 u^{2n} \, e^{-tu^2} \, du = F_n(t)$

related to the error function.

King, Rys, Dupuis (1976): uses orthogonal polynomials to reduce problem to 2-dimensional integrals. Superior for integrals with d and f basis functions.

FLOATING SPHERICAL GAUSSIANS

Simplest basis (subminimal) $\phi = e^{-\alpha(r-r_A)^2}$ for each electron pair, e.g., BeH_2



No SCF needed but much searching needed for big molecules.

MINIMAL-BASIS SETS

Slater type: STO-Old (1930).

STO-NG: give equivalent results.

STO-3G: is chosen for extensive exploration.

This sets up the

HF/STO-3G Model

EXTENDED BASIS SETS

Double zeta: 2 × minimal

Split valence: 4-31G, 6-31G.

Now: 3-21G, 6-21G.

Polarized: $6\text{-}31G^{*}$, $6\text{-}31G^{**}$

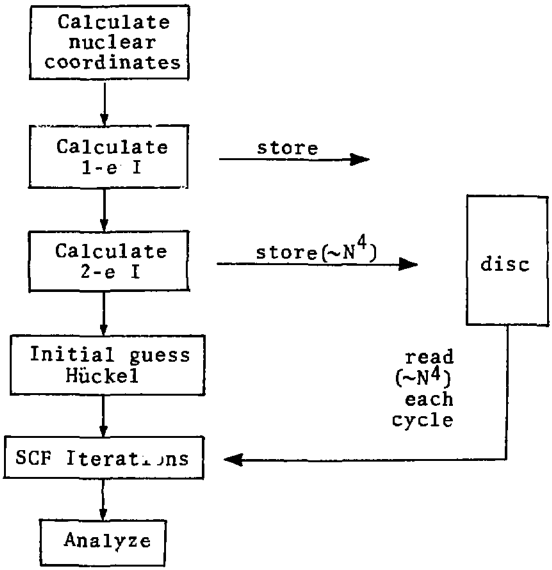
*: d on Li...F

**: also p on H

Large: $6\text{-}31G^{**}$ (suitable for correlation)

also uncontracted (841/41), etc.

STRUCTURE OF SCF PROGRAM



ENERGY DERIVATIVES (HF)

Closed shell:

$$\frac{\partial E}{\partial R} = \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial R} + \tfrac{1}{2} \sum_{\mu\nu\lambda\sigma} [P_{\mu\nu} P_{\lambda\sigma} - \tfrac{1}{4} P_{\mu\lambda} P_{\nu\sigma} - \tfrac{1}{4} P_{\mu\sigma} P_{\nu\lambda}] \frac{\partial(\mu\nu|\lambda\sigma)}{\partial R}$$

$$+ \frac{\partial V^{nuc}}{\partial R} - 2 \sum_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial R} \sum_{i}^{\frac{1}{2}} \varepsilon_i C_{\mu i} C_{\nu i}$$

Integral derivatives needed but do not have to be stored!

HARTREE FOCK TIMES (VAX)

N=40 **15 min**

e.g.,	STO-3G	C_6H_5F
	4-31G	C_3H_6
	6-31G*	CH_3NH_2

N=60 **60 min**

e.g.,	STO-3G	C_8H_{18}
	4-31G	C_5H_8
	6-31G*	CO_3

Derivative calculation requires about the same time as single point → Factor 2.

Optimization

(Fletcher-Powell)

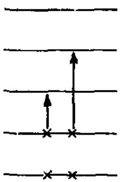
(Murtagh-Sargent)

Approximately one derivative run per variable given a good starting geometry.

MOLLER-PLESSET THEORY

$$\Psi_{MP} = \Psi_{HF} + \tfrac{1}{4} \sum_{ij}^{occ} \sum_{ab}^{vir} a_{ij}^{ab} \, \Psi_{ij}^{ab}$$

Double substitution corrections:



etc.

These are treated as perturbations.

$$H(\lambda) = \sum_p F_p + \lambda[H - \sum_p F_p]$$

where λ = perturbation parameter. If $\lambda=0$, Ψ_{HF} and Ψ_{ij}^{ab} are exact and λ should be 1. Expand in powers of λ and cut off,

$$E(\lambda) = \underbrace{E^{(0)} + \lambda E^{(1)}}_{\text{HF}} + \lambda^2 \underset{\substack{\uparrow \\ \text{used in MP2}}}{E^{(2)}} + \lambda^3 \underset{\substack{\uparrow \\ \text{MP3}}}{E^{(3)}}$$

MP2 THEORY

$$E^{(2)} = -\tfrac{1}{4}\sum_{ij}^{occ}\sum_{ab}^{vir} \frac{|(ij\|ab)|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}$$

$$(ij\|ab) = \chi_i^*(1)\chi_j^*(2)\left(\frac{1}{r_{12}}\right)$$

$$\times\ [\chi_a(1)\chi_b(2) - \chi_b(1)\chi_a(2)]\ d\tau_1 d\tau_2$$

This step requires integral transformation from $(\mu\nu|\lambda\sigma)$ to $(ij\|ab)$. Simple, but $O(nN^4)$ compared with $O(N^4)$ for HF.

SCF PROPERTIES

Lecture 8

by

Ernest R. Davidson

Department of Chemistry
University of Washington
Seattle, Washington

MOLECULAR PROPERTIES FROM *AB INITIO* SCF

I. ENERGETICS
II. CHARGE/SPIN DENSITY
III. "POLARIZABILITIES"

I. ENERGETICS

A. Geometry

1. Isomers
2. Rotomers
3. Reaction intermediates
4. Transition states

Stationary points, $\Delta U = 0$

B. Reaction Energy

1. Barriers
2. Isomerization
3. Rearrangement

$\Delta U = U(\underline{x}_1) - U(\underline{x}_2)$ @ $\Delta U = 0$

C. Normal Mode Analysis, $\underline{K} = (\partial^2 U/\partial x_i \partial x_j)$ @ $\Delta U = 0$

1. Vibrational levels
2. Force constants
3. Normal modes
   a. vibration
   b. reaction

D. Excited States

1. Electronic excitation
2. Ionization potential
3. Electron affinity

Unrelaxed "Koopmans"

$$IP = -\varepsilon_i \qquad \psi \cong \mathrm{Det}|1^1 2^2 \ldots i^1 \ldots|$$

$$EA = -\varepsilon_\alpha \qquad \psi \cong \mathrm{Det}\left|1^1 2^2 \ldots \left(\frac{N}{2}\right)^2 \alpha^1\right|$$

$$\Delta E(^3 i\alpha) = \varepsilon_\alpha - \varepsilon_i - J_{i\alpha}$$

$$\Delta E(^1 i\alpha) = \varepsilon_\alpha - \varepsilon_i - J_{i\alpha} + 2K_{i\alpha}$$

Relaxed orbitals separate SCF on each state

$$\Delta E = U^*(R^*) - U(R)$$

vertical or adiabatic.

II. <u>CHARGE/SPIN DENSITY</u>

$$\rho_c = \rho_\alpha + \rho_\beta = \sum P_{ij}\, f_i f_j$$

$$\rho_s = \frac{(\rho_\alpha - \rho_\beta)}{(N_\alpha - N_\beta)}$$

A. <u>Population Analysis</u>

$$\rho_c \cong \sum \tilde{P}_{iAjB}\, g_{iA} g_{jB}$$

where

g_{iA} = atomic orbital

$\tilde{P}_{iAiB}$ = Bond order $g_{iA} \leftrightarrow g_{jB}$

$$P_{iAiA} + \sum_{jB} P_{iAjB}\, S_{iAjB} = \text{Mulliken population}$$

B. Moments

1. Charge

$$\langle \mathcal{O} \rangle = \sum_\alpha Z_\alpha \mathcal{O}(R_\alpha) - e\int \psi^* \sum_i \mathcal{O}(r_i)\psi \, d\tau_1 \ldots d\tau_N$$

$$\langle \mathcal{O} \rangle = \sum_\alpha Z_\alpha \mathcal{O}(\vec{R}_\alpha) - e\int \mathcal{O}(r) \; \rho(r) d\tau$$

where $\mathcal{O}$ = X,Y,Z dipole moment

$\mathcal{O} = 3X^2-R^2, \quad 3Y^2-R^2, \quad 3Z^2-R^2$
$3XY, \quad 3YZ, \quad 3XZ$ quadrupole moment

$\mathcal{O} = r_A^{-1}$ diagmagnetic shielding (also $\partial U/\partial Z_A$)

$\mathcal{O} = x_A/r_A^3, \quad y_A/r_A^3, \quad z_A/r_A^3$ electric field

Hellmann-Feynman force

$\mathcal{O} = x_A y_A/r_A^5$ etc.

$(3x_A^2 - r_A^2)/r_A^5$ etc.

field gradient "q" (e^2qQ quadrupole coupling)

2. Spin

$$\langle \mathcal{O} \rangle = \int \rho_s(r) \, \mathcal{O}(r) d\tau$$

$\mathcal{O} = \frac{8\pi}{3} \delta(r_A)$ isotropic Fermi contact

$\mathcal{O} = (3x_A^2-r_A^2)/r_A^5$ etc., anisotropic

$$D = \int \psi^* \sum_{i>j} \frac{3(\vec{r}_{ij}\cdot\vec{s}_i)(\vec{r}_{ij}\cdot\vec{s}_j) - \vec{s}_i\cdot\vec{s}_j r_{ij}^2}{r_{ij}^5} \psi d\tau_1 \ldots$$

D,E (spin dipole-dipole part) zero field splitting

3. Derivatives

$\frac{\partial\langle \mu \rangle}{\partial X_A}$ IR intensity, μ = dipole moment

Finite difference

$$\frac{\partial\langle \mu \rangle}{\partial X_A} \cong \frac{\langle \mu \rangle|_{x_A+\delta} - \langle \mu \rangle|_{x_A}}{\delta}$$

4. Vibrational average

$$\langle \mathcal{O} \rangle_n = \int \langle \mathcal{O} \rangle_x \, \psi^2_{\substack{VIB \\ (n)}}(x) \, dx$$

$$\mu_{nn'} = \int \langle \mu \rangle_x \, \psi_{\substack{vib \\ (n)}} \, \psi_{\substack{vib \\ (n')}} \, dx$$

$$\mu_{nn'} \cong \left.\frac{\partial\langle \mu \rangle}{\partial x_A}\right|_e \langle n|x_A|n' \rangle$$

III. POLARIZABILITIES

$$H = H^{o} + \lambda_A A + \lambda_B B + \ldots .$$

$$|\psi^{1}\rangle = \sum_{n} \frac{|n\rangle\langle n|\ \lambda_A A + \ldots\ |0\rangle}{E_o - E_n}$$

$$E = E_o + E_1 + E_2 + \ldots$$

$$E_1 = \text{[illegible]}\ A + \ldots\ |0\rangle$$

$$E_2 = \sum_{n} \frac{|\langle 0|\lambda_A A + \ldots\ |n\rangle|^2}{E_o - E_n}$$

$$A = A_o + A_1 + \ldots$$

$$A_o = \langle 0|A|0\rangle$$

$$A_1 = \sum_{n} \langle 0|A|n\rangle\langle n|\lambda_A A + \ldots|0\rangle + \frac{\langle 0|\lambda_A A \ldots|n\rangle\langle n|A|0\rangle}{E_o - E_n}$$

$$E = E_o + \sum_{A} \lambda_A A_o + \tfrac{1}{2} \sum \lambda_A \lambda_B\ K_{AB}$$

$$\langle A\rangle = A_o + \sum_{B} K_{AB}\ \lambda_B$$

$$K_{AB} = \sum_n \frac{\langle 0|A|n\rangle\langle n|B|0\rangle + \langle 0|B|n\rangle\langle n|A|0\rangle}{E_o - E_n}$$

$$K_{AA} = 2\sum_n \frac{\langle 0|A|n\rangle\langle n|A|0\rangle}{E_o - E_n}$$

$$K_{AB} = \frac{\partial^2 E}{\partial\lambda_A \partial\lambda_B} = \frac{\partial\langle A\rangle}{\partial\lambda_B}$$

One electron operators

$$A = \sum_{i=1}^{N} a(i) \quad , \qquad B = \sum b(i) \ , \qquad \text{etc.}$$

$\alpha\ \beta\ldots$ occupied orbitals

$k\ \ell\ m\ldots$ virtual orbitals

$$K_{AB} \cong \sum_\alpha \sum_k \frac{\langle\alpha|a|k\rangle\langle k|b|\alpha\rangle}{\varepsilon_\alpha - \varepsilon_k}$$

Finite perturbation theory, coupled Hartree-Fock

$$F = F^o(\rho) + \lambda_A A + \lambda_B B + \ldots$$

$$\rho \neq \rho^o$$

$$K_{AB} \cong \frac{\langle A\rangle|_{\lambda_B} - \langle A\rangle_o}{\lambda_B}$$

Example polarizability

$$\lambda_A A = +E_x(-\mu_x)$$

$$\lambda_B B = E_y(-\mu_y) \quad , \text{ etc.}$$

$$\alpha_{xx} = \frac{\partial}{\partial E_x} \langle -\mu_x \rangle$$

$$\alpha_{xy} = \frac{\partial}{\partial E_y} \langle -\mu_x \rangle$$

Polarizability derivatives (Raman intensity)

$$\frac{\partial \alpha_{xx}}{\partial x_A} \quad \text{finite difference}$$

<u>NMR Shielding</u>

$$\text{pert} = \tfrac{1}{2} \sum_j \left(-i\vec{\nabla}_j + \frac{1}{c}\vec{A}'_j\right)^2 + \tfrac{1}{2}\sum_j \nabla_j^2$$

$$A'_j = \underbrace{\tfrac{1}{2}\vec{H}_o \times \vec{r}_j}_{A_j} + \sum_M g_M \beta_M \frac{\vec{I}_M \times \vec{r}_{jM}}{r^3_{jM}}$$

Gauge invariant atomic orbitals

$$e^{-i\vec{A}_M \cdot \vec{r}/c} \, g_{kM}$$

<u>NMR Coupling</u>

$$
\text{pert} = \sum_{M,k} \left\{ \frac{2\beta\hbar}{i}\, g_M\beta_M \frac{\vec{I}_M \cdot \vec{\ell}_{kM}}{r_{kM}^3} \right.
$$

$$
+ \; g\beta\hbar\, g_M\beta_M \frac{3(\vec{s}_k\cdot\vec{r}_{kM})(\vec{I}_M\cdot\vec{r}_{kM}) - r_{kM}^2\, \vec{S}_k\cdot\vec{I}_M}{r_{kM}^5}
$$

$$
\left. + \; \underbrace{g_\beta g_M\beta_M\hbar\, \frac{g\pi}{3}\, \delta(\vec{r}_{kM})\, \vec{S}_k\cdot\vec{I}_M}_{\text{dominant term}} \right\}
$$

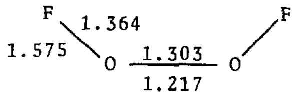
BOND LENGTHS

$\mathrm{C} \overset{1.101\ \text{Å}}{\underset{1.128}{\text{———}}} \mathrm{O}$ (calc) (expt)

$\mathrm{F} \overset{1.33}{\underset{1.44}{\text{———}}} \mathrm{F}$

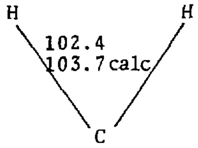
$\mathrm{N} \overset{1.06}{\underset{1.10}{\text{———}}} \mathrm{N}$

$\mathrm{O} \overset{1.160}{\underset{1.207}{\text{———}}} \mathrm{O}$

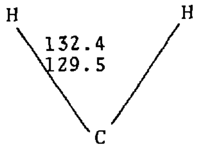


Typical AB - 0.04 Å

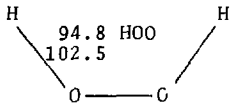
AH - 0.01 Å



^3B



^1A



111.5° HOOH
113.7°

typical ±2°

ENERGY

DISSOCIATION ENERGY

CO	11.1 eV	7.9
F_2	+1.35 eV (expt)	-1.37 calc
O_2	5.08 eV	1.3
N_2	9.9 eV	5.3

ISODESMIC (retention of bond type) error 5 kcal/mol

$CO_2 + CH_4 \rightarrow 2H_2CO$ 52.2 kcal (calc)

57.9 (expt)

Hydrogenation

$H_2O + 2H_2 \rightarrow CH_4 + H_2O$ -63.5 calc

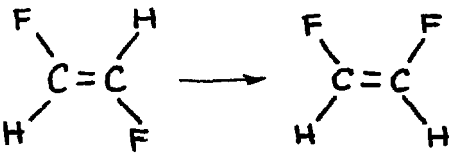
-57.3 expt

Hydrogen Transfer

$2CH_4 + C_2H_4 \rightarrow 2C_2H_6$ -13.0 calc

-17.2 expt

ISOMERIZATION ENERGY



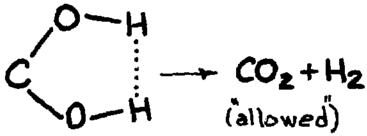
ΔE calc - 0.26 kcal

expt - 1.2

Requires near HF limit basis

ENERGY BARRIERS

C_2H_6 rotation	3 ± ½ kcal
Rotation (in general)	±0.5 kcal SCF limit
H_2 + H exchange	9.8 kcal (exact)
	24.4 kcal SCF



81	kcal SCF
68	kcal CI

IONIZATION AND EXCITATION ENERGIES

Formamide		Koop	Expt	I.P.
$HCONH_2$	n	11.9	10.3	
	π_1	11.5	10.5	
	π_2	15.6	14.2	
	σ	16.5	14.8	

Urea $(NH_2)_2CO$

	unrelaxed	ΔSCF	CI (or expt)
$^3n\pi^*$	7.6	5.9	6.8
$3\pi\pi^*$	7.6	6.0	6.7
n ion	11.2	8.4	9.1
π ion	10.6	8.6	9.4
$^1n\pi^*$	7.9	6.1	--
$^1\pi\pi^*$	11.2	10.1	(7.1)
Formamide			
$^3n\pi^*$	6.2	4.5	(5.3)
$^3\pi\pi$	6.2	5.2	--
$^1n\pi$	6.7	4.8	(5.1)
$^1\pi\pi^*$	9.8	8.2	(7.3)

FORCE CONSTANTS

LiH	1.10 (calc)	1.03 (expt)	
N_2	26.1	23.0	
CH_4	sym. stretch	5.7 (calc)	5.5 (expt)
	t stretch	5.5	5.4
	e bend	0.64	0.58
	t bend	0.61	0.54

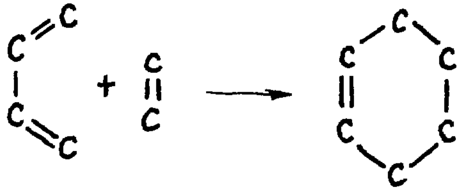
H_2	$\frac{\partial^2 E}{\partial R^2}$	0.3814 calc SCF	0.3701 exact
	$\frac{\partial^3 E}{\partial R^3}$	-1.2686	-1.2703
	$\frac{\partial^4 E}{\partial R^4}$	4.309	4.224

H_2CO	CO	13.66	12.90
	CO-CH	0.79	0.74

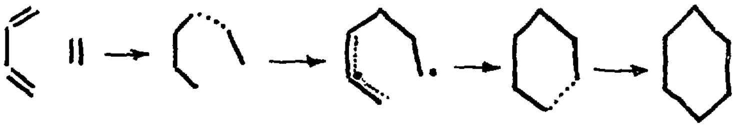
TRANSITION STATES $\Delta U = 0$

One negative force constant. No experimental data?

Controversy



Symmetrical or



DIPOLE MOMENT (DEBYE)

	SCF	CI	Expt'l
LiH	6.002	5.853	5.82
BeH	0.282	0.248	--
BH	-1.733	-1.470	--
CH	-1.570	-1.427	-1.40
NH	-1.627	-1.587	--
OH	-1.780	-1.633	-1.66
FH	-1.942	-1.816	-1.82
$\frac{\partial\mu}{\partial R}$ LiH	0.23	0.30	0.29
$\frac{\partial\mu}{\partial R}$ HF	1.7	--	0.95

		SCF		Expt
LiF	q	-11.8		-0.128
CO	Q	-1.81		-1.63

OSCILLATOR STRENGTH

C_2H_4 $\quad {}^1\pi\pi^*$ $\quad$ SCF 0.113

Expt 0.34

		Transition	Moment
H_2	$\sigma_g^2 \to \sigma_g\pi_\alpha$	0.367 (SCF)	0.360 (expt)
	$\to \sigma_g\sigma_\alpha$	0.55 (SCF)	0.50 (expt)

Large molecule canonical orbital usually meaningless.

IVO has 50% error, ΔSCF difficult; has non-orthogonality problem.

FERMI CONTACT

SCF 0 Expt ?

Planar π radicals

BO	$(^2\Sigma)$	$\|\psi(0)\|^2$		Aℓ O	
	SCF	Expt		SCF	Expt
B	0.70	0.72	Aℓ	0.00	0.7
O	0.05	0.02	O	0.08	--

Anisotropic

B	0.34	-0.06
O	0.19	0 *or* 0.8

Na atom, Q = 217G (SCF), 316G (expt)

	CH_2 ^3B SCF	$\|\psi(0)\|^2$ expt
C	0.214	0.22
H	0.007	~0

		expt
D	0.763	0.76
E	0.062	0.052

STATIC POLARIZABILITY

N_2	$\alpha_\parallel$	$\alpha_\perp$	(a_0^3)
calc	14.79	9.75	
expt	15.02	10.32	

H_2S	calc	3.47	expt.	3.67	mean α $Å^3$
HF $\bar{\alpha}$	calc	4.98	expt	5.60	
$\alpha_\parallel - \alpha_\perp$		1.45		1.49	
H_2O	calc	8.68	expt	9.82	
CO	calc	2.43	expt	1.95	

J NMR COUPLING

CF in CH_3F	calc	-98
	expt	-162
HF	calc	784
	expt	530

Coupled SCF (FPT)

HH in CH_4	-6.1 calc	-12.4 expt
Geometry HH in C_2H_4	12.9	2.5
cis in C_2H_4	6.0	11.6
trans in C_2H_4	14.9	19.1

Sum over "state"

HH in CH_4	-58	coupled SCF FPT	
HH in H_2	60 calc	43 expt	FPT

Very sensitive to basis set choice.

CH in CH_4 319 calc , 125 expt
(minimum basis FPT)

NOTE: FPT-INDO works well, *ab initio* unreliable!

Shielding Constants ^{13}C

	calc	expt
C_2H_6	- 7.4	- 8.0
CH_3F	-65.4	-77.5
C_2H_4	-130.8	-125.6
H_2CO	-199.6	-197

GENERALIZED VALENCE BOND

Lecture 9/10

by

William A. Goddard, III

Department of Chemistry
California Institute of Technology
Pasadena, California

THEORETICAL OBJECTIVES: *DIFFERENCES* IN ENERGIES

e.g.,
- Bond energy
- Barrier height
- Excitation energy
- Ionization potential
- Electron affinity
- Potential surface

A small bias toward either limit causes a big change in the ΔE. Theory must treat all states at comparable levels of:

- Basis set
- Orbitals
- Level of CI

REVIEW – ELECTRONIC STATES

$$\mathcal{H}\Psi = E\Psi$$

$$\mathcal{H}(1,2,\ldots N) = \underbrace{\sum_e h_r + \sum_{e>e'} \frac{1}{r_{ee'}}}_{\mathcal{H}^{e\ell}} + \sum_{M>M'} \frac{Z_M Z_{M'}}{R_{MM'}} + \sum_M h_M$$

where M = nuclei, e = electrons

$$h_e = -\tfrac{1}{2}\nabla^2 - \sum_M \frac{Z_M}{r_M}$$

BORN-OPPENHEIMER APPROXIMATION

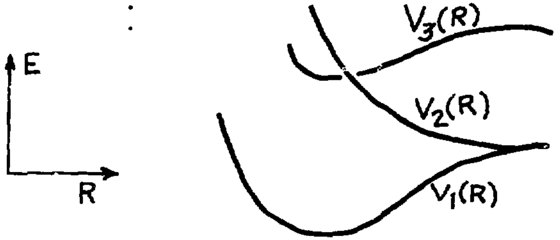
$$\Psi(r_e\ldots,R_M\ldots) = \underbrace{\psi^{e\ell}(r_e\ldots)}_{\substack{\text{electronic}\\ \text{wave function}}}\ \underbrace{F^{NUC}(R_M\ldots)}_{\substack{\text{vibration}\\ \text{rotation}}}$$

Solve for $\psi^{e\ell}$ as a function of geometry

$$\mathcal{H}^{e\ell}\ \psi_1^{e\ell}(r_e\ldots) \quad = \quad E_1^e\psi_1^{e\ell}(r_e\ldots) \qquad \text{ground state}$$

$$\mathcal{H}^{e\ell}\ \psi_2^{e\ell} \quad = \quad E_2^{e\ell}\psi_2^{e\ell} \qquad \text{first excited state}$$

$$\vdots$$



WAVEFUNCTIONS

$$\psi^{e\ell}(r_1, r_2, \ldots r_N)$$

HARTREE-FOCK APPROXIMATIONS

(Also molecular orbital)

$$\psi(1\ldots N) \quad = \quad \mathcal{A}\{(\phi_1\alpha)(\phi_1\beta)(\phi_2\alpha)(\phi_2\beta)\ldots\}$$

where $\mathcal{A}$ is a Slater determinant, ϕ_i = molecular orbital, and α,β = up and down spin. Apply variation principle; get HF or SCF equation, $H^{HF}\ \phi_i = \varepsilon_i\phi_i$. This is a one-electron equation but gets N eigenstates.

$$H^{HF} \quad = \quad h + \sum_j (2J_j - K_j)$$

where the $2J_j$ is the Coulomb operator and K_j the exchange operator. The H^{HF} depends on occupied orbitals, therefore solve iteratively.

General Advantages of HF

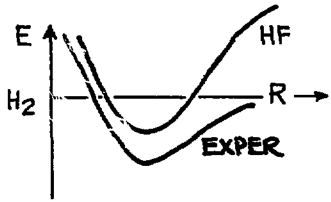
1. Orbital interpretation – useful for qualitative reasoning.

2. Good for qualitative interpretation of PES (photo-electron spectroscopy).

General Problems of HF

1. Does not describe bond breaking or reactive intermediates.

e.g.,



2. Bad for excitation energies,

   a. HF O_3 is a triplet, not a singlet

   b. Systematic bias against d^M

Ni	Exper	HF
d^{10}	1.8 eV	5.6 eV
s^1d^9	-0.04 eV	1.8 eV
s^2d^8	0	0

3. Basic problem with HF (weakly overlapping radical orbitals), i.e., broken bond,

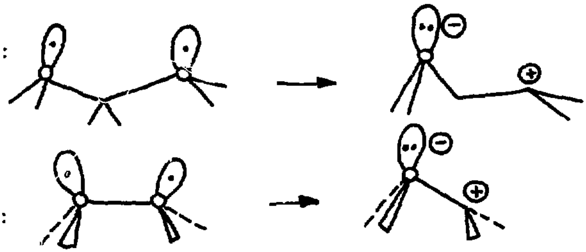
$$\phi_g(1)\phi_g(2) = (\chi_\ell(1)+\chi_r(1))(\chi_\ell(2)+\chi_r(2))$$

$$\underbrace{(\chi_\ell\chi_\ell + \chi_r\chi_r)}_{\text{ionic}} + \underbrace{(\chi_\ell\chi_r + \chi_r\chi_\ell)}_{\text{covalent}}$$

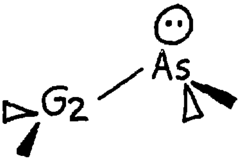
therefore *force* ionic character. How does system respond (closed shell case)?

$S_A(111)$:

$S_A(100)$:



Guideline: When can MO theory be trusted? (band theory, tight binding, EHT) — When chemical ideas would lead to doubly occupied orbitals. Therefore don't trust $S_A(111)$, $S_A(100)$, or $S_A(110)$, but reconstructed GaAs(110) may be OK.



Hartree-Fock Wavefunction (Molecular Orbital)

Bond energies (D_e)

	Theory	Experiment	
CH_3-CH_3	72.1	96.7	kcal
CH_3-OH	62.9	98.8	
HO-OH	1.0	52.2	
$H_2C{=}CH_2$	123.3	180.3	
$H_2C{=}O$	105.5	182.1	

(Good basis, DZd)

Conclusion: HF not useful for bond energies

Approximate versions HF

1. Extended Hückel theory, tight binding CNDO, MINDO, MNDO. Semiempirical parameterized to fit one property or another.
2. X_α use $\rho^{1/3}$ approximation to exchange terms. Muffin tin approximation not semiempirical.
3. Pseudopotential approximation to replace core orbitals.

Advantages of Approximate HF

* Good geometries for simple (closed shell) molecules
* Simple prediction of photoemission (using Koopman's theorem)

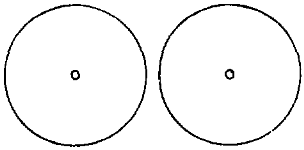
Disadvantages

* Even exact HF theory has serious deficiencies for our purposes.

SW-xα (Scattered Wave)

Approximate κ as $\alpha\rho^{1/3}$, where ρ = electron density and α = parameter (~0.7).

Muffin tin approximation:



H^{HF} is spherically symmetric within sphere about each nucleus and constant between spheres. The advantage is no atomic basis set (but do need scattered wave basis). Problems are:

a. Do not get total energy, therefore cannot get geometries and potential surfaces (use of $\sum_i \varepsilon_i$ leads to linear H_2O; O not bound to Ni surface).

b. Bad PES unless muffins overlap (violates theory).

c. Spherical averaging bad if atom not symmetric.

Semiempirical

Use minimal basis (one function per AO). Evaluate $\langle \chi_\mu | H^{HF} | \chi_\nu \rangle$ semiempirically. Do not get total energy, therefore there is a problem to get the geometry.

For the extended Hückel theory (EHT), put average two-electron terms into one electron part,

$$\alpha = \langle \chi_\mu | H^{HF} | \chi_\mu \rangle \qquad \text{same atom}$$

$$\beta < \langle \chi_\mu | H^{HF} | \chi_\nu \rangle \qquad \text{adjacent atoms}$$

For iterative EHT (SCCC), put charge term into α.

For CNDO/2 (INDO), evaluate largest J_{ij} (atomic κ_{ij}) from theory and get α and β semiempirically (to fit theory).

For MINDO, MINDO/2, MINDO/2.5, MINDO/3, MNDO: it is the same as CNDO but choose parameters to fit experiment.

All have serious problems with transition metals since there is not enough experimental data to fix all parameters. All are bad for reaction intermediates. MINDO systematically bad for closed vs. open.

Second Problem with Approximate HF

$$E_{total} = E_0 + E_1 + E_2$$

where E_0 is the nuclear-nuclear, E_1 = electron-nuclear, and E_2 = electron-electron

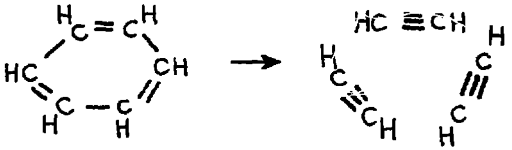
$$\sum_i \varepsilon_i = E_1 + 2E_2$$

the ε_i are HF one-electron energies (Koopman's IP), and $2E_2$

double counts electron-electron; therefore,

$$E_{total} = \sum_i \varepsilon_i + (E_0 - E_2)$$

Many methods calculate only ε_i, therefore, cannot get E_{total}. (Most pseudopotential calc., most tight bonding, therefore, cannot get geometric structure. Example:

EHT ⇒ 

without a barrier. Experiment: benzene more stable than 3 HCCH by ~4 eV and these are large barriers in both directions.

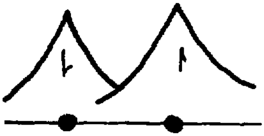
GENERALIZED VALENCE BOND (GVB)

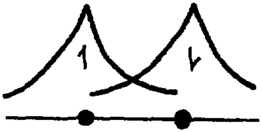
Solve for orbitals while *including dominant electron correlation* effects. Basic wavefunction has *one orbital per electron* (not two electrons in orbital), but orbitals allowed to overlap. Normal bond pair qualitatively similar to valence bond wavefunction. Two orbitals, one on each atom. *Ab initio* (no adjustable parameters)

CORRELATED WAVEFUNCTIONS

HF 

two electrons in same orbital

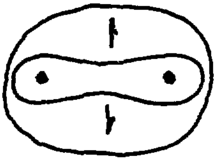


GVB 

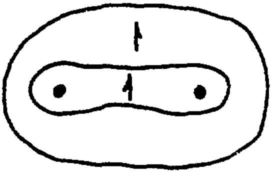
Correlate motion of electrons along axis (left-right correlation)

ADDITIONAL CORRELATION EFFECTS

i) (two cases):
(call this up-down or starboard-portside or angular correlation)



ii) (call this in-out or tight-loose correlation)



Generally *four important correlations* include all four in GVB calculation (5 orbitals to describe one electron pair), denote as (1/5)

BEST SIMPLE WAVEFUNCTION

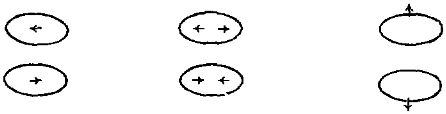
$$
\begin{aligned}
\psi(1,2) \;=\; \Big\{ & C_1\phi_{\sigma g}(1)\phi_{\sigma g}(2) - C_2\phi_{\sigma M}(1)\phi_{\sigma M}(2) && \text{left-right} \\
- & C_3\phi_{2\sigma g}(1)\,\phi_{2\sigma g}(2) && \text{in-out} \\
- & C_4\phi_{\pi ux}(1)\phi_{\pi ux}(2) - C_4\phi_{\pi uv}(1)\phi_{\pi uy}(2) \Big\} && \text{angular}
\end{aligned}
$$

Error = 2 kcal at R_e or 0 kcal at $R=\infty$. This is a (1/5) calculation.

CORRELATION EFFECTS (kcal) WITHIN A BOND PAIR

Elect.1

Elect.2

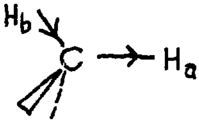


H_3C-CH_3	8.7	2.0	1.0+1.0
H_3C-OH	12.8	1.6	0.9+0.8
HO-OH	29.8	1.1	0.6+0.5
H_2C=CH_2 [σ]	5.8	1.8	1.1
[π]	17.3	1.3	0.4
H_2C=O [σ]	8.8	1.6	1.0
[π]	23.0	1.1	0.5

All these correlation errors disappear at $R=\infty$.

Conclusion: DO GVB (1/5) calculation on bond being dissociated. Four intra-pair correlations account for 50% of HF bond energy error.

ADDITIONAL CORRELATIONS



When the electrons in bond CH_a move toward H_2 then the electrons in bond CH_b move toward the C. But at $R=\infty$ this correlation disappears.

These interpair correlations generally increase the bond energy.

Cross Correlation Effects

(a) Double Bond ($> = <$)
σ bond pair ($> \rightarrow <$)
π bond pair ($> \leftarrow <$)

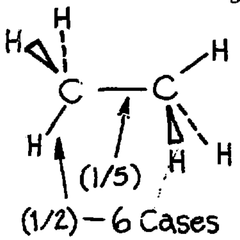
Correlated motion: $H_2C = CH_2$, 6.8 kcal; $H_2C = O$, 9.3 kcal.

(b) Adjacent bonds

Simultaneous correlations

H_3C-CH_3	CC-CH	1.2 kcal	*6	=	7.2 kcal
H_2C=CH_2	CCσ-CH	1.0 kcal	*4	=	4.0
	CCπ-CH	1.2	*4	=	4.8
H_2C=O	COσ-CH	0.7	*2	=	1.4
	-OM	3.4	*1	=	3.4
	COπ-CH	0.7	*2	=	1.4
	-OM	4.0	*1	=	4.0

Typical GVB-CI: CC bond energy H_3C-CH_3



thus GVB is (7/17) (7 is the number of electron pairs, 17 is number NO). After optimizing all orbitals do CI (quadruples). Within GVB space (17 orbitals), same calculation for fragments.

<u>Generalized Valence Bond GVB-CI</u>

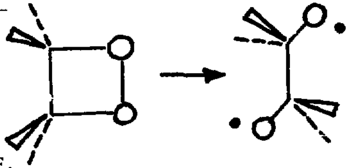
Bond energies (D_e)

	Theory	Exper
CH_3-CH_3	93.9	96.7 kcal
CH_3-OH	98.0	98.8
HO-OH	56.4	52.2
H_2C=CH_2	171.6	180.3
H_2C=O	174.6	182.1
O=O('Δ)	98.2	97.6

(Good basis, DZd)

Conclusion: GVB-CI satisfactory.

Typical calculation

Ring opening: 

Basis: DZd = 68 B.F.

GVB: (i) Include all four correlations for bond being broken (O-O), therefore, 5 orbitals/1 electron pair = 1/5

(ii) Correlate all other valence pairs as in normal GVB (2 orbitals/1 electron pair = 1/2)

Four CH

One CC

Two CO

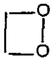
Four O lone pair ,

therefore, $5 + 2 \times 11 = 27$

optimum GVB valence orbitals.

Think of this as 12 occupied MO's plus 15 optimal correlating orbitals. GVB-CI: do high order CI (quadruple excitations) among GVB orbitals (impossible for full basis).

Result: D_{OO} = 14 kcal for ring opening; therefore,

24 kcal strain energy in 

(exper. 26 for  and )

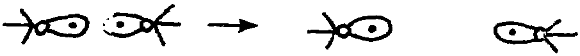
All accurate methods involve configuration interaction.

HF-CI:

1. Calculate set of optimum occupied orbitals
2. Select set of unoccupied (virtual orbitals)
3. Allow single, double, triple... excitations from occupied to virtual orbitals (usually double excitations)

Comments:

a. No reliable method of using less than all virtuals plus all occupied valence orbitals
b. Often do excitations WRT one configuration. This is biased against state with large correlation error.



small correlation large correlation

c. AB + C → A + BC (planar)

Doubles WRT one dominant ~40000 configuration.

CI (HF + S + D)

Include all configurations involving single and double excitations

$$\psi = \psi_o + \sum_{Ii} C_{Ii}\, \psi_i^I + \sum_{\substack{I,J\\ i,j}} C_{IJij}\, \psi_{ij}^{IJ}$$

1. No reliable method of using less than all virtual orbitals and all valence occupied orbitals, therefore, magnitude of calculation increases rapidly with size of system.

2. Inconsistent if number of electrons or bondedness changes. For example, He + He: at $R=\infty$ we need S+D on left He and S+D on right He, therefore, for He_2 we must use S+D+T+Q. If we do only S+D for He_2, do not go to proper He limit at $R=\infty$.

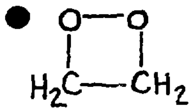
CH_4: at $R=\infty$ CH_3 + H requires doubles on CH_3; therefore, require selected triples in CH_4. If we do all triples on CH_4, this does not lead to proper CH_3 at $R=\infty$.

GVB-CI

1. Calculate orbitals self-consistently while including dominant electron correlation effects and generalized valence bond.
2. Do high order CI (e.g., quadruple excitations) among GVB orbitals and low order CI involving virtual orbitals.
3. GVB orbitals localize into bond orbitals in different regions. Thus we can identify active GVB orbitals for high correlation and inactive orbitals for medium correlation.

   Active: change in process (bond pair being broken)

   Inactive: Not changing



Example: HF-CI

68 basis FNS (DZd)

16 molecular orbitals (12 valence, 4 core)

52 virtual (unoccupied) orbitals

Singly excited determinants, 1.2×10^3

Doubly excited determinants, 3.9×10^5

Triply excited determinants, 2.0×10^8

Quadruply excited determinants, 2.7×10^{10}

(Need at least triples for bond energy; practical level is $\sim 3 \times 10^5$.)

Example: GVB-CI

Optimize orbitals with dominant correlations present. Therefore, 4 core + 24 valence (12 valence in HF).

Residual correlation energy

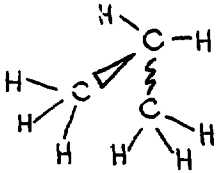
Quadruple excitations within GVB; pairs → 1221 configurations. Cross terms + excitations to virtuals: 8000 to 15000 configurations. Includes major effect of 2.7×10^{10} configs.

Correlation-consistent CI (CC-CI)

Active orbitals: changed directly in the physical process being studied, the CH_a bond pair in our case.

Demiactive orbitals: localized adjacent to the active orbitals and hence responsible for differential correlation effects. The other three CH bonds in CH_4.

Inactive orbitals: other orbitals.

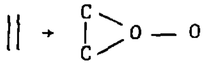


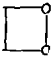
CC-CI: All double excitations out of active orbitals times all single excitations out of demiactive orbitals. Thus, CC-CI includes selected triple excitations, but does not include all doubles. CC-CI increases very slowly with increasing substituents.

$$CH_4 \rightarrow CH_3 + H$$

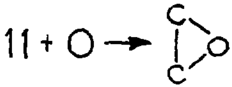
	D_e	Number config.	Error
HF	88.3 kcal	1	1.03 eV
CI(HF+S+D)	99.5	769	0.47
GVB	106.4	5	0.215
CC-CI	111.8	1033	0.016
Exper	112.2		

<u>Problems with GVB</u>: Must use finite complex, therefore, most useful for cases with localized interactions. Can't calculate modification of bulk band structure; is tedious to get photoemission. No reliable semiempirical versions.

Dewar MINDO/3 concludes that $^1O_2 + {}^1O_2 +$ 

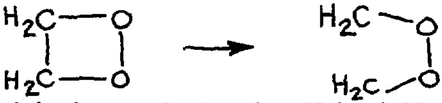
 $\rightarrow$ $^1(H_2CO)^* + H_2O$ OH = 45 kcal

therefore concludes that decomposition must require S $\leftrightarrow$ T intersystem crossing. The problem is that MINDO is biased toward ring geometry by ~20 to 30 kcal., e.g.,

e.g.  OH is 22 kcal too negative

e.g.  D=45 kcal, therefore, MINDO $\Rightarrow$ negative strain of +7 kcal

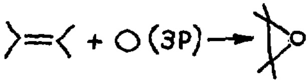
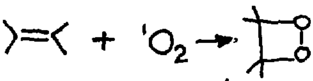
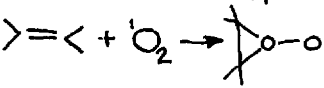
Hinze CNDO-MCSCF + empirical E correction concludes

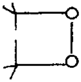


Problem with those calculated $\rightarrow$ CC bond 58, experiment 89.

$\rightarrow$ OO bond 153, experiment 53.

Comparison GVB-CI vs. MINDO/3 ΔH (kcal)

	GVB-CI	MINDO/3	Exper.
	-79.8	-105.3	-83.
	-36.9	-65.5	
	+16.5	-16.1	

 $\rightarrow$ $2H_2CO$ (S_0+S_0) GVB-CI, -50.2 ; MINDO/3, -27.5

$2H_2CO$ (S_0+T_1) GVB-CI, +21.9 ; MINDO/3, +44.4

BASIS SETS

MBS (minimal basis set)

One FN per atomic orbital, therefore 5 on C or O, 1 on H (e.g., STO-3G)

VDZ or DZ (valence double zeta)

Two FNS per atomic orbital (allows contraction upon bond formation), therefore, 9 on C or O, 2 on H.

DZd: Add d FNS on C or O,

Add p FNS on H (if break CH bond)

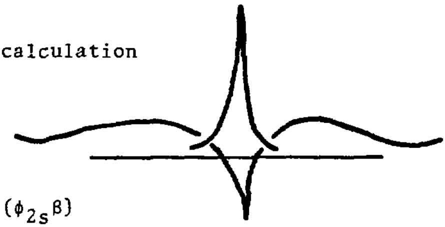
This allows polarization of bond orbitals

$$\phi = \sum_{\mu} C_{\mu} \chi_{\mu}$$

where ϕ = GVB orbital, μ = sum over all centers, and χ = basis functions.

Effective Potentials, replace Ar core of Ni with effective potential V_{core}. Therefore, reduce system to 10 electrons. Get V_{core} from all-electron *ab initio* (Hartree-Fock) calculation on several states (6) of atom. Leads to *ab initio* result for molecules. (V_{core} completely determined by atomic calculation, requires energies and shapes to be reproduced by V_{core}.)

Comments: pseudopotential calculation



Be $(1s)^2(2s)^2$

HF: $\mathcal{A}\underbrace{(\phi_{1s}\alpha)(\phi_{1s}\beta)}_{\text{core}}\ \underbrace{(\phi_{2s}\alpha)(\phi_{2s}\beta)}_{\text{valence}}$

$$H_N\phi_{2s} = \varepsilon_{2s}\phi_{2s}$$

$$H_N = h + \underbrace{V_{core}}_{2J_{1s}-K_{1s}} + \underbrace{V_{val}}_{2J_{2s}-K_{2s}}$$

where J_{1s} is Coulomb and K_{1s} is exchange energy

Pseudopotential: find V_{core} to replace core electrons.

Pseudopotentials.

Note, let $\quad \phi_V = \phi_{2s} + \lambda\phi_{1s}$

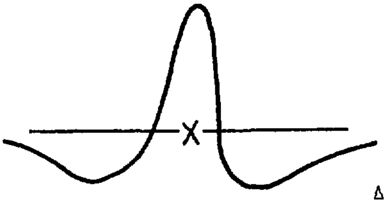
$$\mathcal{A}(\phi_{1s}\alpha)(\phi_{1s}\beta)(\phi_{2s}\alpha)(\phi_{2s}\beta) = \mathcal{A}(\phi_{1s}\alpha)(\phi_{1s}\beta)(\phi_V\alpha)(\phi_V\beta)$$

therefore, can mix ϕ_{1s} into ϕ_{2s} to get new ϕ_V without changing energy or properties. Usually choose $\lambda \ni \phi_V$ is smooth.

<u>Problem</u>: consider density of

orthogonal orbitals: $\rho = \underset{\text{core}}{2\phi_{1s}^2} + \underset{\text{valence}}{2\phi_{2s}^2}$

Nonorthogonal orbitals: $\langle \phi_N | \phi_{1s} \rangle = S \neq 0$



$$\rho = \underset{\text{core}}{2\phi_{1s}} + \underset{\text{valence}}{2\phi_V^2} + \Delta\rho$$

$$\Delta\rho = -\frac{4s}{1-s^2}\,\phi_{1s}(\phi_N - s\phi_{1s})$$

$\Delta\rho$ = + core, - valence

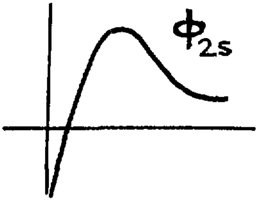
Pseudopotential calculations usually assume $\rho_{val} = 2\phi_V^2$ (summed over valence). This is wrong, it leads to charge too small in core region and too large in bond region.

<u>Another problem</u>: valence-valence interaction

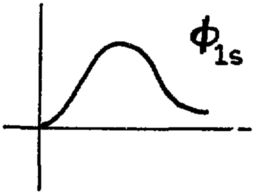
$$\phi_{2s} \qquad \phi_N = \phi_{2s} - \lambda\phi_{1s}$$

changes valence. Example: $J_{2s,2s} \neq J_{V,V}$.

The correct HF Hamiltonian involves ϕ_{1s} in complicated way.



self energy

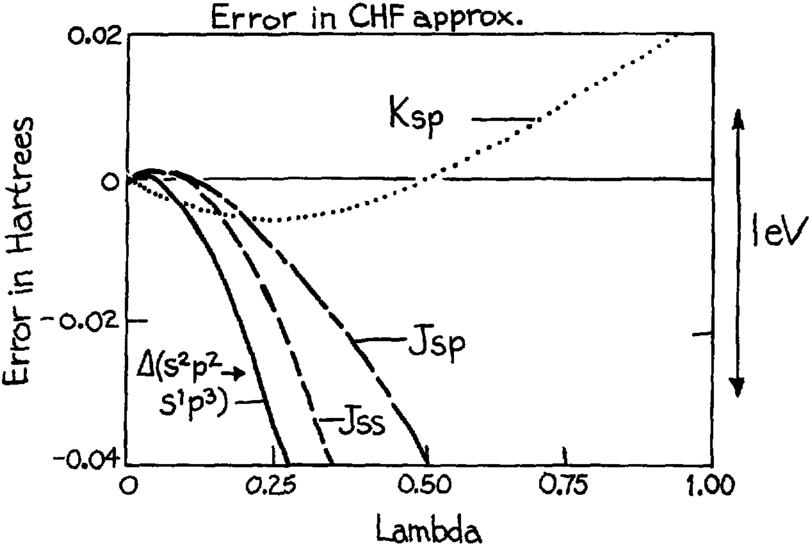


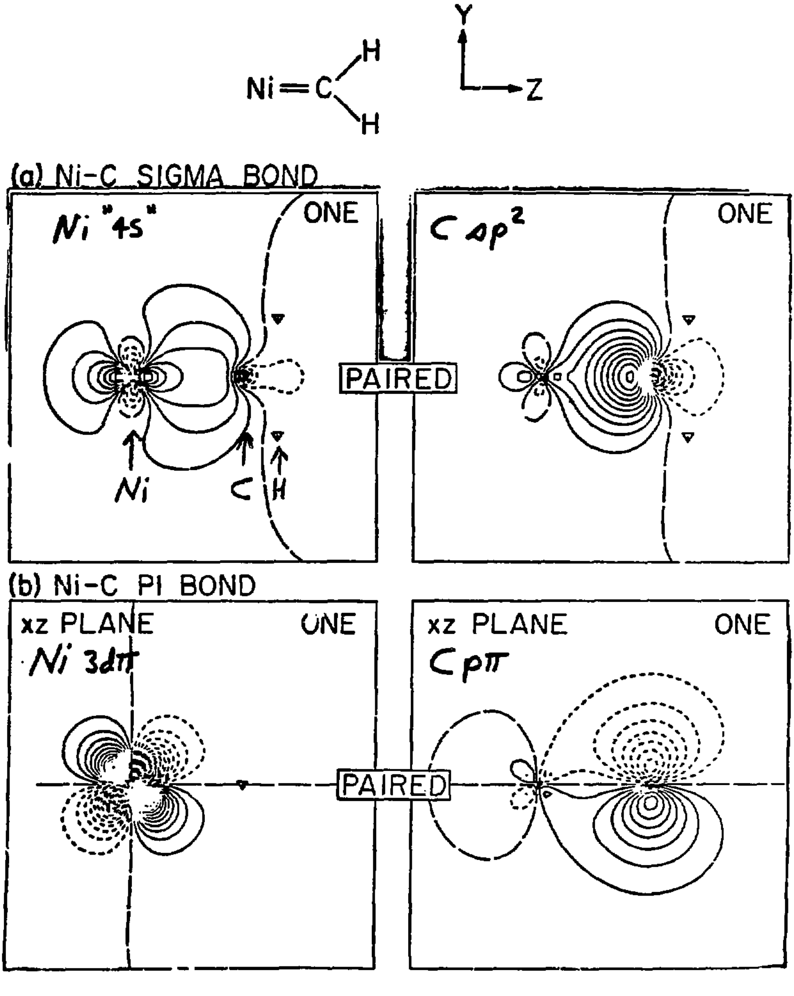
different

$$|2s'\rangle = |2s\rangle + \lambda\,|1s\rangle$$

HF Orbitals

Pseudo-orbital

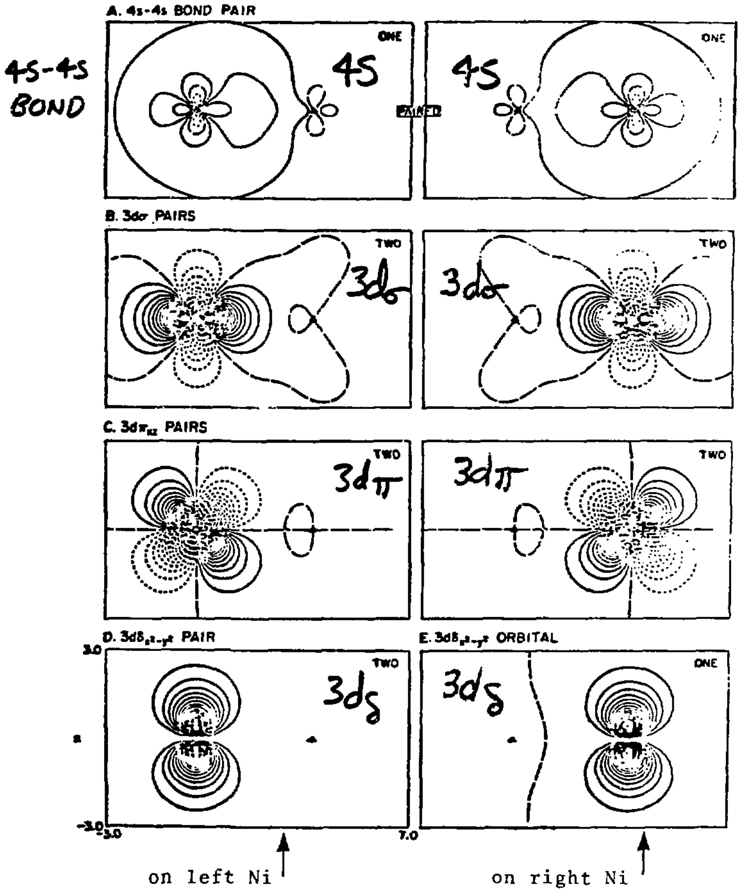



Y
Z
Ni
C
H
H
(a) Ni-C SIGMA BOND
Ni "4s"
ONE
Ni
C
H
PAIRED
C sp²
ONE
(b) Ni-C PI BOND
xz PLANE
Ni 3dπ
ONE
PAIRED
xz PLANE
C pπ
ONE

Ni_2 Qualitative picture:

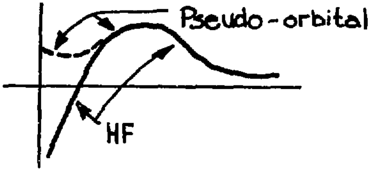
$$\underset{\text{left}}{(4s)^1(3d)^9} \;—\; \underset{\text{right}}{(4s)^1(3d)^9}$$

Ni_2 GVB ORBITALS (δδ STATE)



AB INITIO EFFECTIVE POTENTIALS

1. Require shape of pseudo orbital to be unchanged in valence region.



2. Require that valence-valence interactions be unchnaged

a. $J_{2s,2s} = J_{\overline{2s},\overline{2s}}$, $K_{2s,2p} = K_{\overline{2s},\overline{2p}}$

(ACE) $J_{2s,2p} = J_{\overline{2s},\overline{2p}}$,

b. Require the combination entering the valence

(SHC) Hamiltonian to be correct

3. Find smooth core shape to satisfy above conditions.

After choosing smooth pseudo-orbitals, find effective potential for core, such that

$$H_N \phi_{\bar{N}} = \varepsilon_N \phi_{\bar{N}} \qquad (\varepsilon_N \text{ from ab initio})$$

n

where

$$H_N = h + V_{core} + V_{val}$$

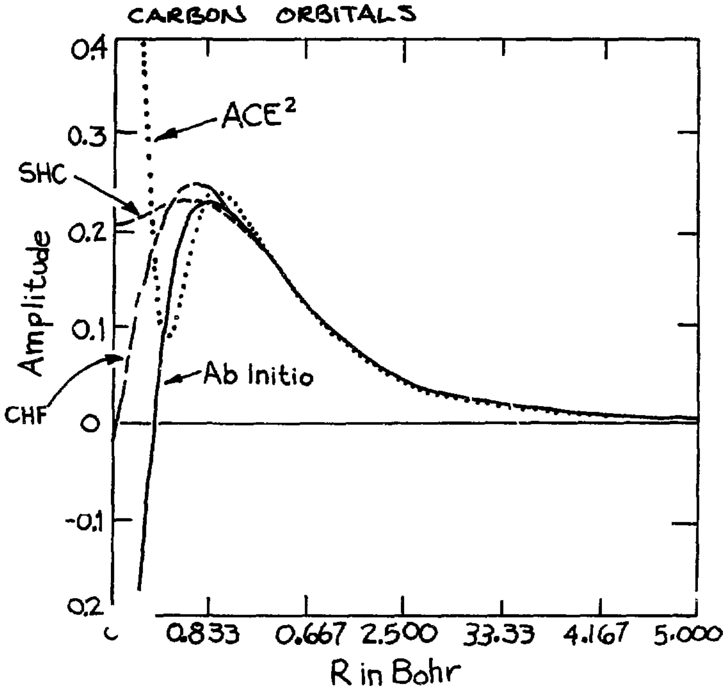
(h is one-electron, V_{val} is valence Hamiltonian. Do this for lower states, e.g.,

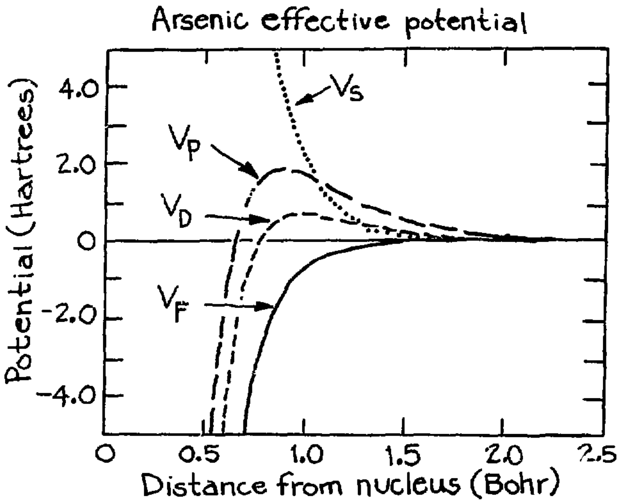
$$\text{Ga: (a)} \underbrace{(1s)^2(2s)^2(2p)^6(3s)^2(3p)^6(3d)^{10}}_{\text{cores}}(4s)^2(4p)^1$$

(b) (core)$(4s)^1(4p)^2$

(c) (core)$(4s)^2(4d)^1$

(d) (core)$(4s)^2(4f)^1$


CARBON ORBITALS
0.4
0.3
0.2
0.1
0
-0.1
0.2
ACE²
SHC
CHF
Ab Initio
Amplitude
0.833
0.667
2.500
33.33
4.167
5.000
R in Bohr


Arsenic effective potential
Potential (Hartrees)
4.0
2.0
0
-2.0
-4.0
Vs
Vp
VD
VF
0
0.5
1.0
1.5
2.0
2.5
Distance from nucleus (Bohr)

Example, Ga

$$V_{Ga}^{core}(r) = V_s(r)\hat{P}_s + V_p\hat{P}_p + V_d\hat{P}_d + V(r)$$

where $\hat{P}_s$ = projection operator for $\ell=0$, WRT Ga center

$\hat{P}_p$ = projection operator for $\ell=1$

$\hat{P}_d$ = projection operator for $\ell=2$

V_s, V_p, V_d = repulsive (Pauli principle)

$$V(r) = -\frac{3}{r} - \frac{1.136}{r} e^{-2.715r^2} - 1.429\, e^{-0.965r^2}$$

$$V_s(r) = \frac{13.119}{r^2} e^{-1.884r^2} + 7.042\, e^{-1.858r^2}$$

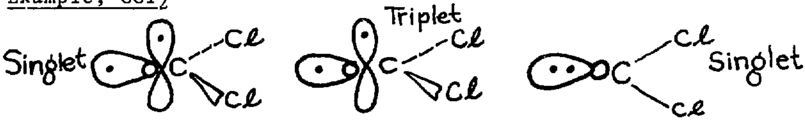
$$V_p(r) = \frac{5.106}{r^2} e^{-0.449r^2} - 0.950\, e^{-0.451r^2}$$

$$V_d(r) = 0.906\, e^{-0.445r^2}$$

States of atoms: error $\leq$ 0.01 eV

States of molecules: error $\leq$ 0.1 eV

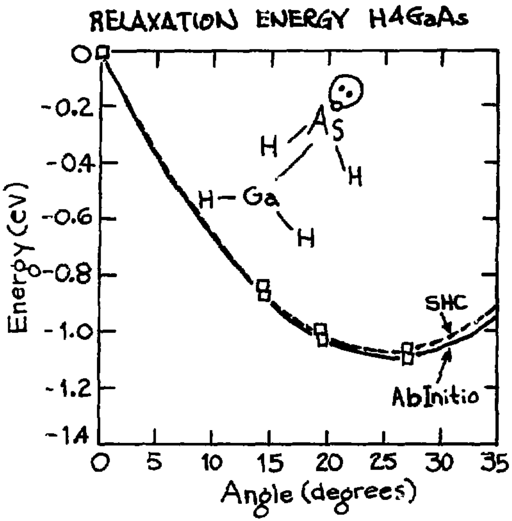
Example, CCl_2



CCl_2	ab initio	effective potential	error
singlet	2.233 eV	2.238 eV	+0.005 eV
triplet	0.820 eV	0.824 eV	+0.004 eV
singlet	0	0	

EXAMPLE: TIMING

Ab initio	26 min	IBM 360/91
Effective potential	26 sec	IBM 3032

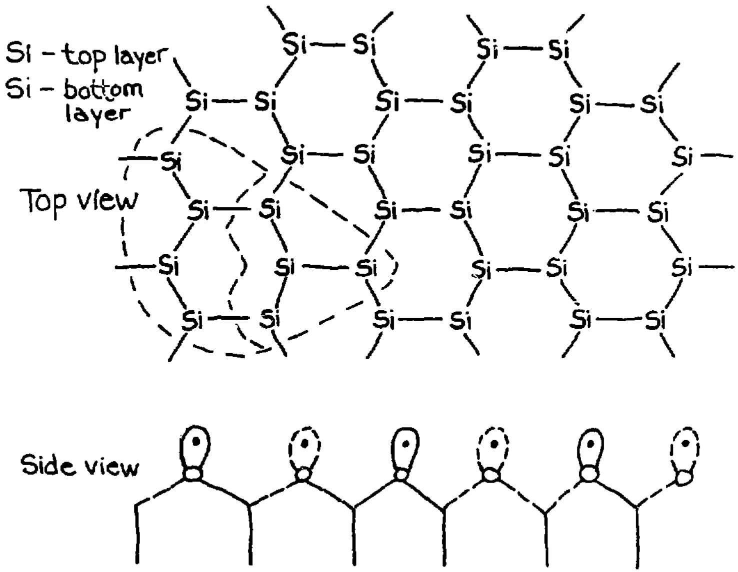


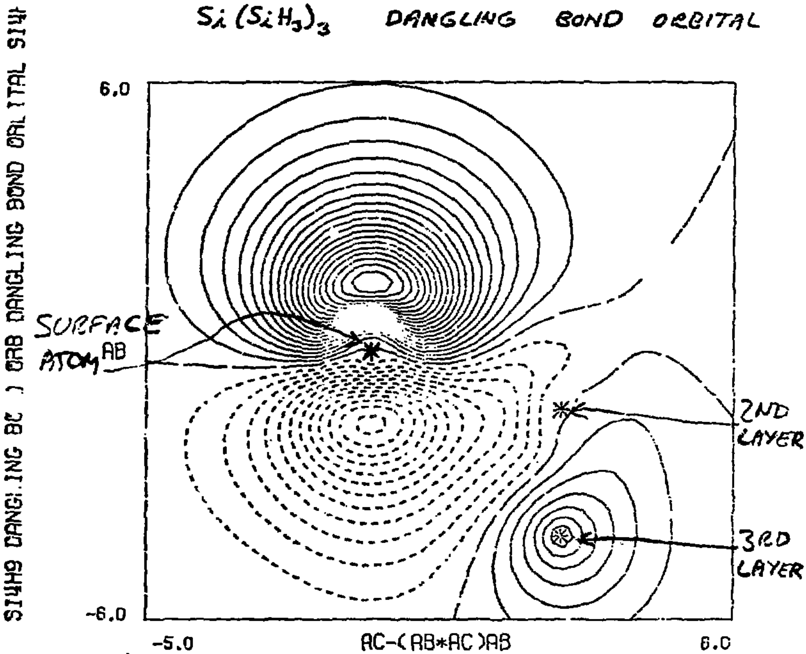
	ab initio	Eff. Pot.	Error	
Ω	25.927	25.559	0.368	Twist angle (deg)
E_{min}(cV)	-1.093	-1.0727	0.02 eV	Relaxation energy

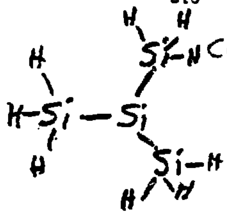
Example: Si atom $(3s)^2(3p)^2$

hybrid $(3s)^1(3p)^3$

Four singly occupied orbitals, therefore four bonds (tetrahedral). Crystalline Si: four bonds to each Si (tetrahedral). Si surface: three bonds to each surface Si, therefore one electron in hybrid orbital pointing away from surface (dangling bond)





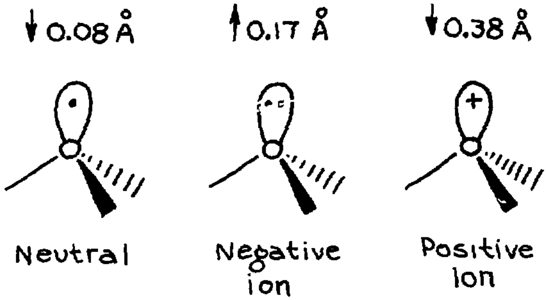


CONTOURS: $\Delta = 0.01$, $R = -0.12$ TO $+0.19$

ZOPT

RELAXATION OF SURFACE Si

THEORY

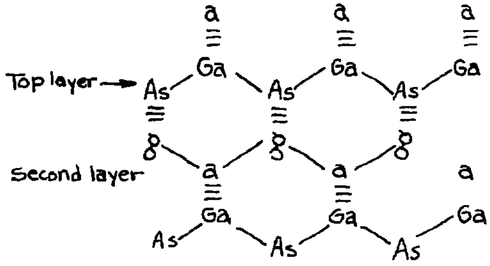


EXPERIMENT (after theory was published)

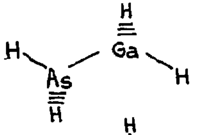
Neutral	Neutral
	2 × 1 (Mönsch)
0.12 ± 0.04 Å	One Si ↓0.16 Å
	One Si ↓0.00 Å
(1 × 1 stabilized with Te)	Average displacement ↓ 0.08 Å

GaAs RECONSTRUCTION AND INITIAL STEPS IN OXIDATION

IDEAL SURFACE

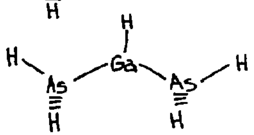


Model Ga_1As_1:



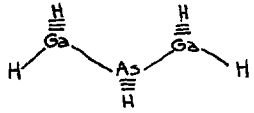
$Ga_1As_1H_4$

Model Ga_1As_2:



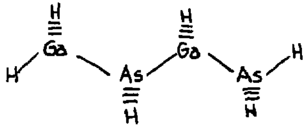
$Ga_1As_2H_5$

Model Ga_2As_1:



$Ga_2As_6H_5$

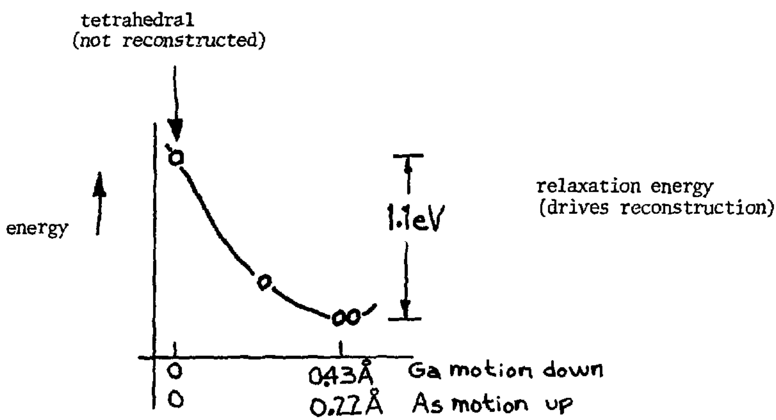
Model Ga_2As_2



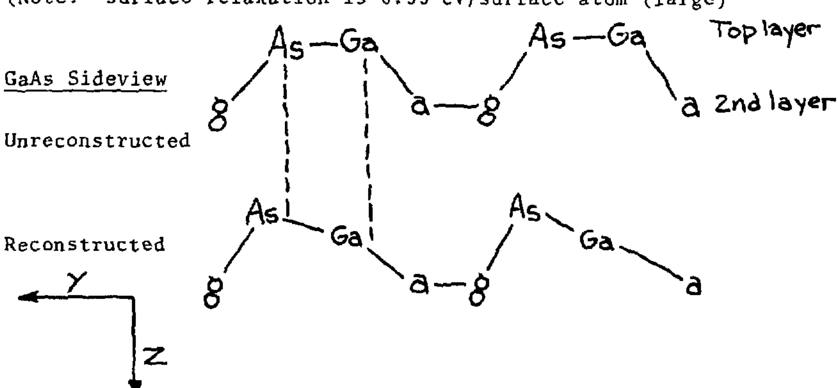
$Ga_2As_2H_6$

GEOMETRY VARIATIONS

Fix second layer. Move all As of first layer same. Move all Ga of first layer same. H: use standard AsH, GaH distance. Orientation determined by virtual position of Ga or As which is represented by H.

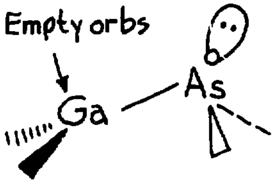


(Note: surface relaxation is 0.55 eV/surface atom (large)



Model	δz_{Ga}	δz_{As}	δy_{Ga}	δy_{As}	Twist angle
Ga_1As_1	0.43	-0.22	0.48	0.37	25.6°
Ga_1As_2	0.41				
Ga_2As_1	0.44				
Exper.	0.43 to 0.47	-0.20 to -0.23			27 ± 2°

Conclusion: local model describes essence of reconstruction in GaAs. (In progress, GaP, AℓAs, AℓP.)



CHARACTER OF WAVE FUNCTIONS

Geometry = angles at surface atom

As: 90, 90, 108 → 96° average

Ga: 125, 125, 108 → 119° average

Isolated trivalent molecule: 93° AsH_3, 120° GaH_3.

Conclusion: surface reconstruction dominated by local valence effects.

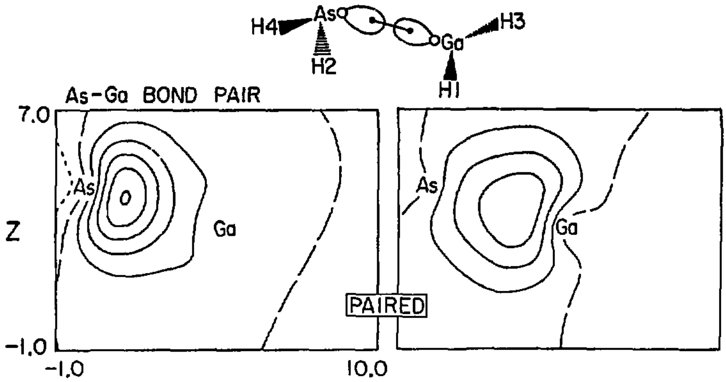
CHARGES

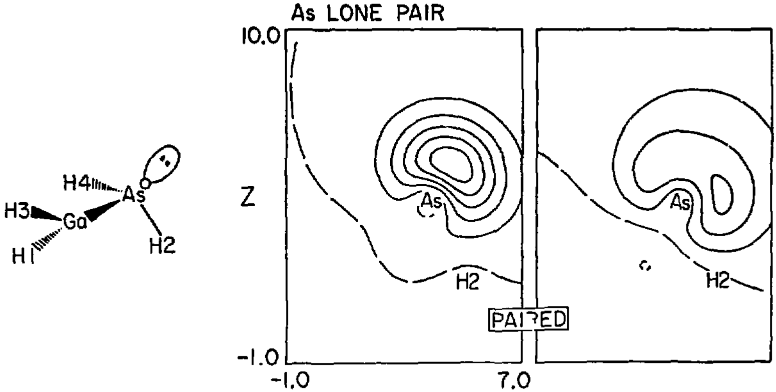
Ga-As bond pair, 0.3 electrons from Ga to As (Mulliken population)

Ga: 48% P character, As: 84% P character.

As lone pair, localized on As, 33% P character.

Note: geometry of surface determined by local valence effects. Therefore, can estimate effects from experiment or theoretical study of small complexes.


H4
As
H2
Ga
H3
H1
As-Ga BOND PAIR
7.0
-1.0
Z
As
Ga
PAIRED
-1.0
10.0

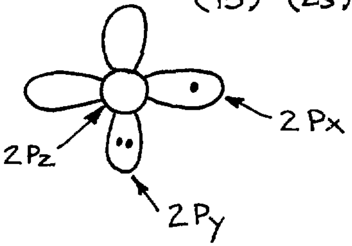


H4
As
H3
Ga
H1
H2
As LONE PAIR
10.0
Z
As
H2
PAIRED
-1.0
-1.0
7.0
As
H2

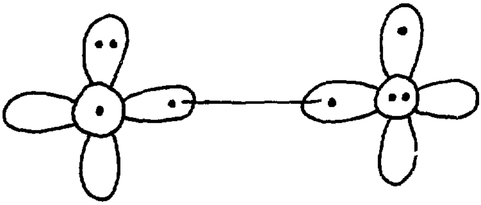
OXIDATION

O ATOM

$(1s)^2 (2s)^2 \underbrace{(2P)^4}$



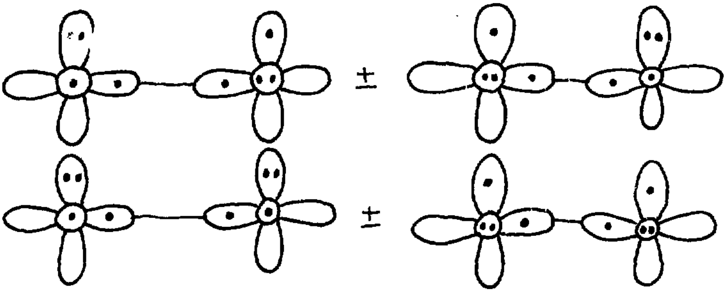
O_2 MOLECULE



+ Resonant configuration

biradical — triplet lowest since orbitals orthogonal

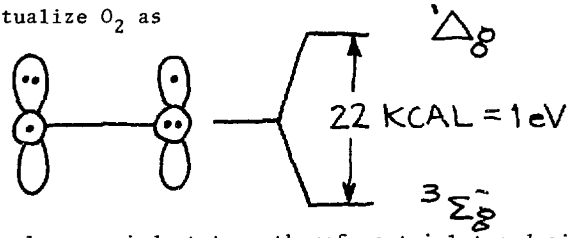
∴ triplet
and singlet



Ground State $^3\Sigma_g$

Excited states

Conceptualize O_2 as



two singly occupied states, therefore triplet and singlet. These two orbitals orthogonal, therefore triplet lower.

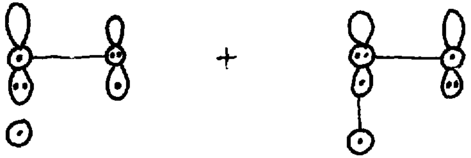
BOND ENERGIES

$^3O_2$ σ bond → 47 kcal

π bond → 71 kcal (includes resonance)

$^1O_2$ π bond 22.5 kcal weaker

Bond H to $^3O_2$: Normal H-O bond = 104 kcal (H-OMe or H-OEt)

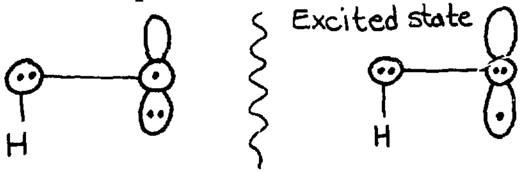


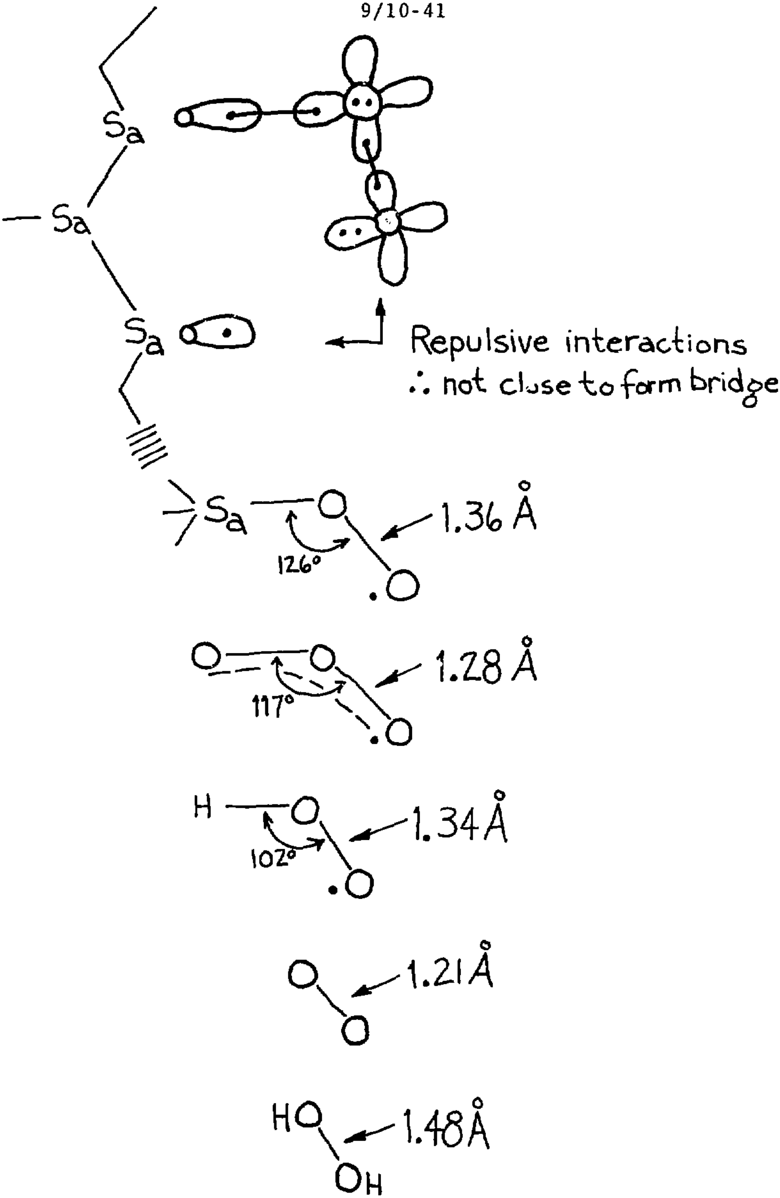
no bond　　　good bond

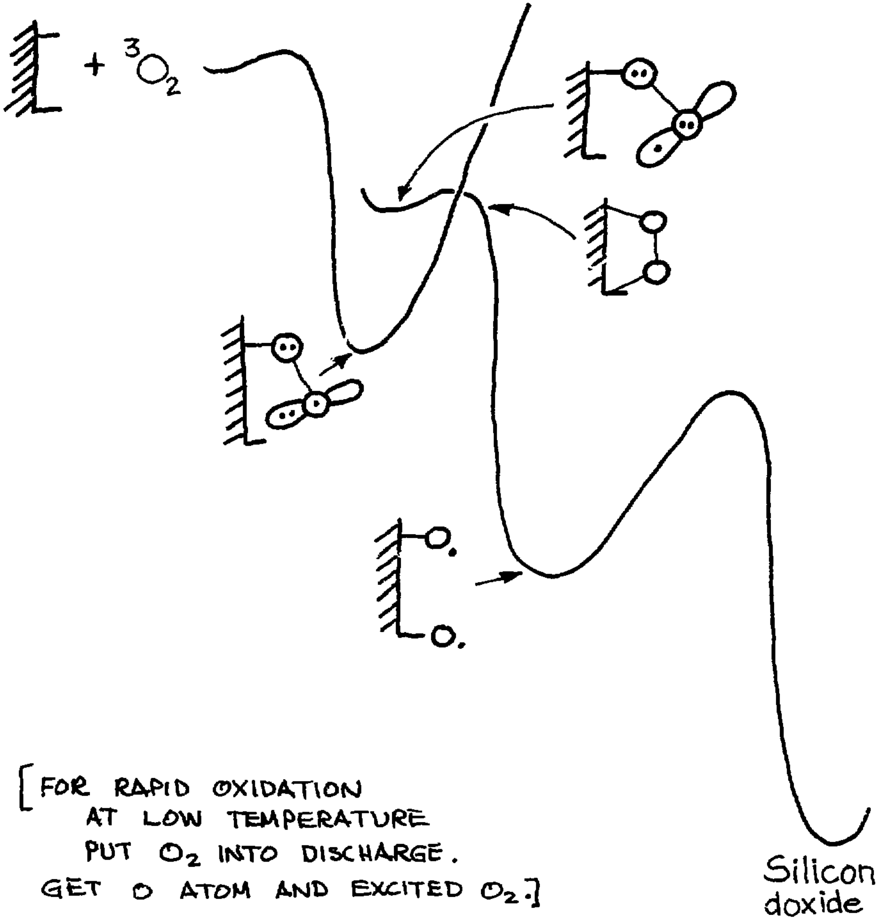
Result: lose most (57 kcal) of resonance); therefore

$D(H\text{-}O_2) = 104 - 57 = 47$ kcal

Bond H to HO_2: Lose remaining 14 kcal (71-57) of O_2 bond; therefore, $D(H\text{-}O_2H) = 104 - 14 = 90$ kcal. Visualize HO_2 as

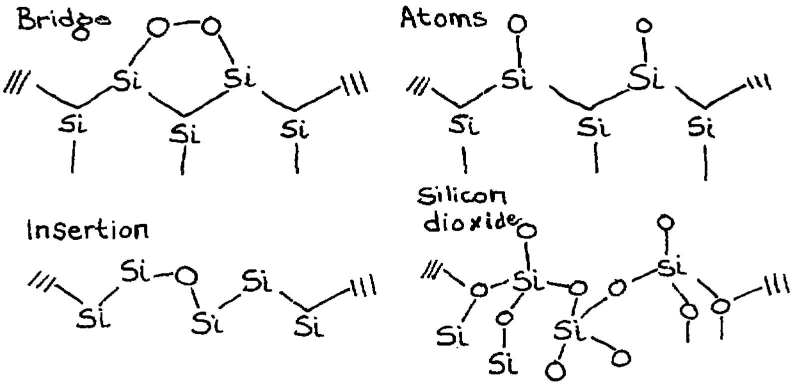



Sa
Sa
Sa
Repulsive interactions
∴ not close to form bridge
Sa
126°
1.36 Å
117°
1.28 Å
H
102°
1.34 Å
1.21 Å
HO
OH
1.48 Å

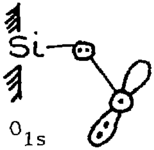


[FOR RAPID OXIDATION
AT LOW TEMPERATURE
PUT O_2 INTO DISCHARGE.
GET O ATOM AND EXCITED O_2.]

Add O_2 to Si surface, and what do we get?

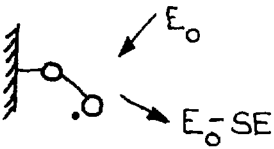


Experimentally it is possible to form one or two intermediate surface oxides before obtaining fully oxidized surface.



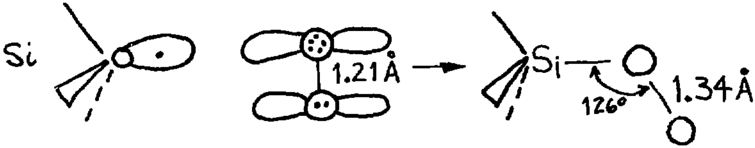
EXPERIMENTAL EVIDENCE FOR PEROXY RADICAL MODEL

1. chemisorbed O_2 on Si.
   a. XPS (Rowe): initial state chemisorbed O_2, two O_{1s} signals of same intensity but different chemical shift
   b. High resolution electron energy loss.
      3 vibrational modes with dipole
      (component) — to surface
      Peroxy SiO stretch
      SiOO bend
      OO stretch



Bond O_2 to As

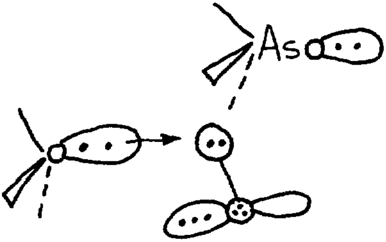
Ground state O_2 requires singly occupied orbital



but as orbital is doubly occupied,

GaAs

forces O_2 to bond

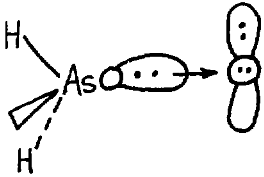


1. Energy is 1.6 eV higher than free $^1O_2$, therefore, no bond.
2. Chemical shift of O_{1s} orbitals is 3.4 eV (disagrees with experiment).

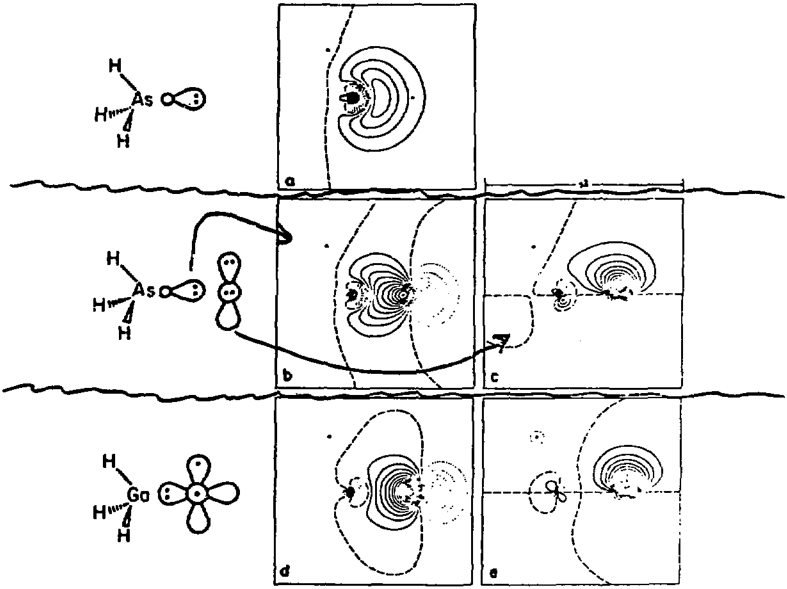
Conclusion: we do not get a chemical bond of O_2 at As

loose complex (bond 0.2 to 0.4 eV)

no significant effect on surface



Basis	Wavefn.	R_{AsO}	Bond energy
DZ	HF	1.74Å	-0.55 eV
DZd	HF	1.61Å	0.64 eV
DZd	GVB-CI	1.63Å	2.25 eV


H
As
H
H
a
H
As
H
H
b
c
H
Ga
H
H
d
e

	$Me_3As = 0$ Exper.	$H_3As = 0$ Theory
R_{AsO}	1.631 Å	1.63 Å
Dipole moment	1.14 eÅ	1.00 eÅ
AsO vibration frequency	107 MeV 112 MeV	124 MeV
Bond energy	--	2.25 eV

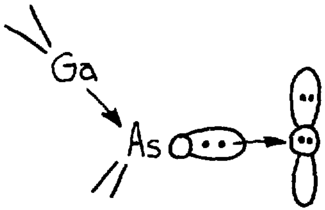
Comparisons of bond lengths

	x-0	x=0	change
P	1.62 Å	1.39 Å	0.23 Å
As	1.80	1.63	0.17
Sb	2.0		
	↑ from X_2O_3	↑ from X_2O_5	

<u>CHEMICAL SHIFTS IN CORE ORBITALS Ga(3d), As(3d)</u>

Upon reconstruction, As(3d) ↓ 0.24 eV (deeper)
Ga(3d) ↑ 0.20 eV
As (lone pair) ↓ 0.86 eV
Ga (empty) ↑ 1.22 eV

Upon oxidation, As(3d) ↓ 2.6 eV
Ga(3d) ↓ 0.8 eV
Ga(empty) ↓ 0.8 eV



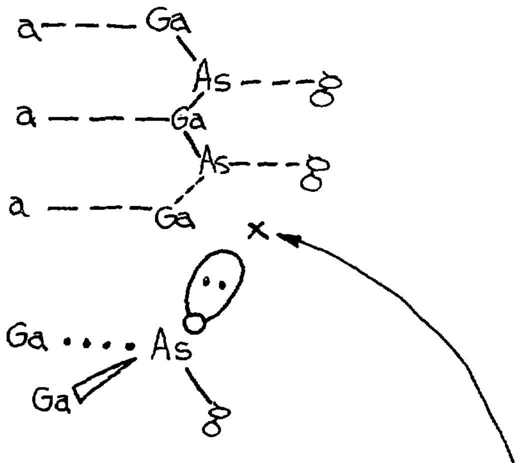
<u>Experimental</u>	Low Coverage Ga	Low Coverage As	High Coverage Ga	High Coverage As
Spicer (1978)	0	+2.9	1.0	4.6
Brundle (1979)	+0.8	+2.8		
Spicer (1979)	+1.0	+2.9		
Donor acceptor	+0.8	+2.6		

Conclusion: experimental chemical shifts of low coverage oxide are consistent with donor-acceptor complex.

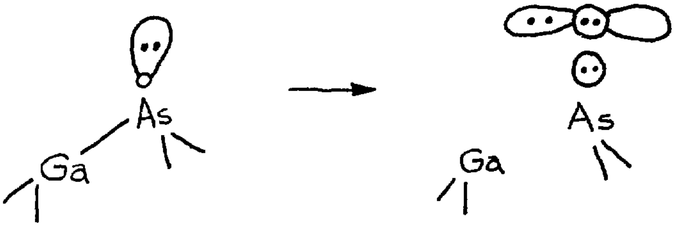
<u>Al Overlayer on GaAs</u>

top view

side view



General assumption is that additional metal (eg., Al) is at normal Ga site.



twist 25.6°   18.6°

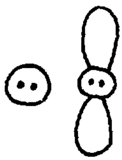
As=O bond is at angle of 56° WRT surface normal

As=O bond = 1.63 Å

EXAFS January 1979, 1.52 Å and May 1979, 1.62 ± 0.1 Å

CHARGE DISTRIBUTION

H_3As  $\mu_{\text{lone pair}}$ = 1.26 eÅ, therefore lone pair centered 0.63 Å from As



$\mu_{\text{lone pair}}$ = 2.32 eÅ

$\mu_{\text{oxy}\pi}$ = -0.26 eÅ

net change, 0.80 eÅ and R = 1.63 Å, therefore, one-half electron transferrred.

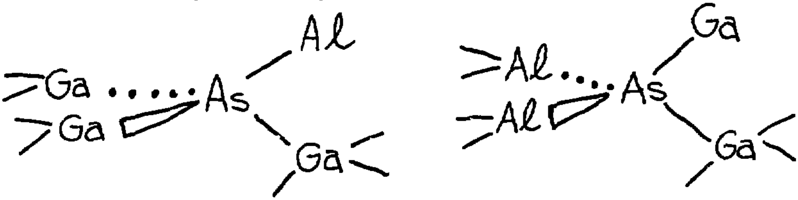
Brillson: one-half monolayer Al on GaAs(110).

Obs: ordered 1 × 1, chemical shifts: ↑ density at Ga

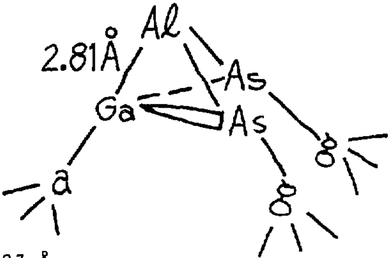
↓ density at Al

no change at As

Mele: tight binding calculation - two cases



Finds that this (on the right) leads to expected change in density. Therefore, concludes that one-half monolayer of Al on GaAs(110) leads to Ga overlayer and Al in surface layer.



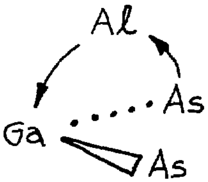
Ab Initio Calculation

Optimum structure

(bond energy 0.56 eV)(HF)

Comparisons, H_3Ga-Al, R = 2.23 Å

H_3As-Al, R = 3.45 Å



Charge transfer

therefore chemical

shifts Al ↓ (deeper)

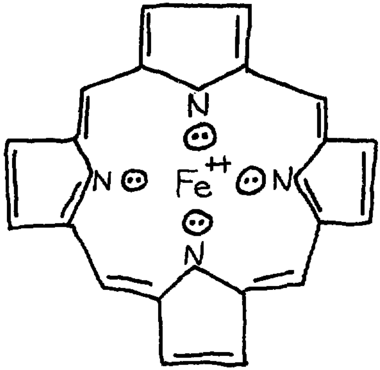
Ga ↑ 0.95 eV

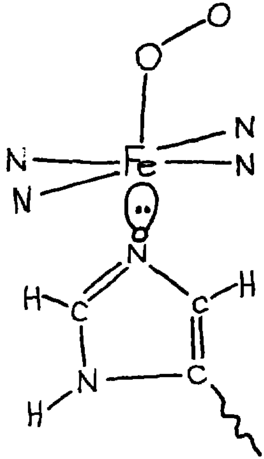
As ↓

Reconstruction: H_5As_2Ga, δZ_{Ga} = 0.41 Å

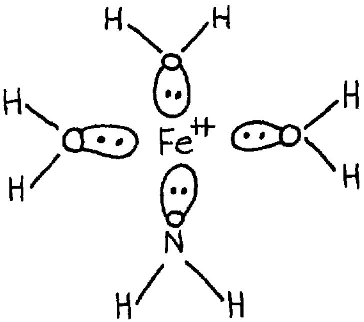
H As GaAl, δZ_{Ga} = 0.29 Å

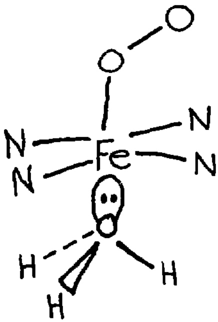
EXAMPLE: Heme Fe





Model Calc

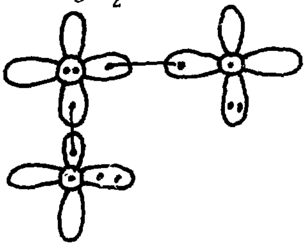




Before bonding O_2 to Fe consider bonding O_2 to O (making ozone)

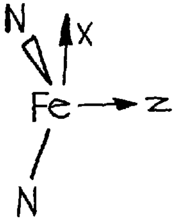
O_2 →

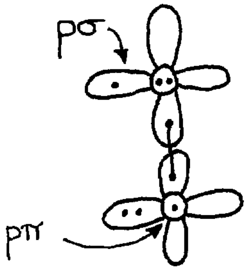
O →



We gain new O-O σ bond (4-7 kcal) and lose O_2 resonance (-57 kcal). Net bond is -10 kcal. Add in new π bonding (+35 kcal), and net bond is +25 kcal. This π bond is special 3-center-4-electron bond (essential to stability of ozone).

Bond O_2 to Mb





Fe q state O_2

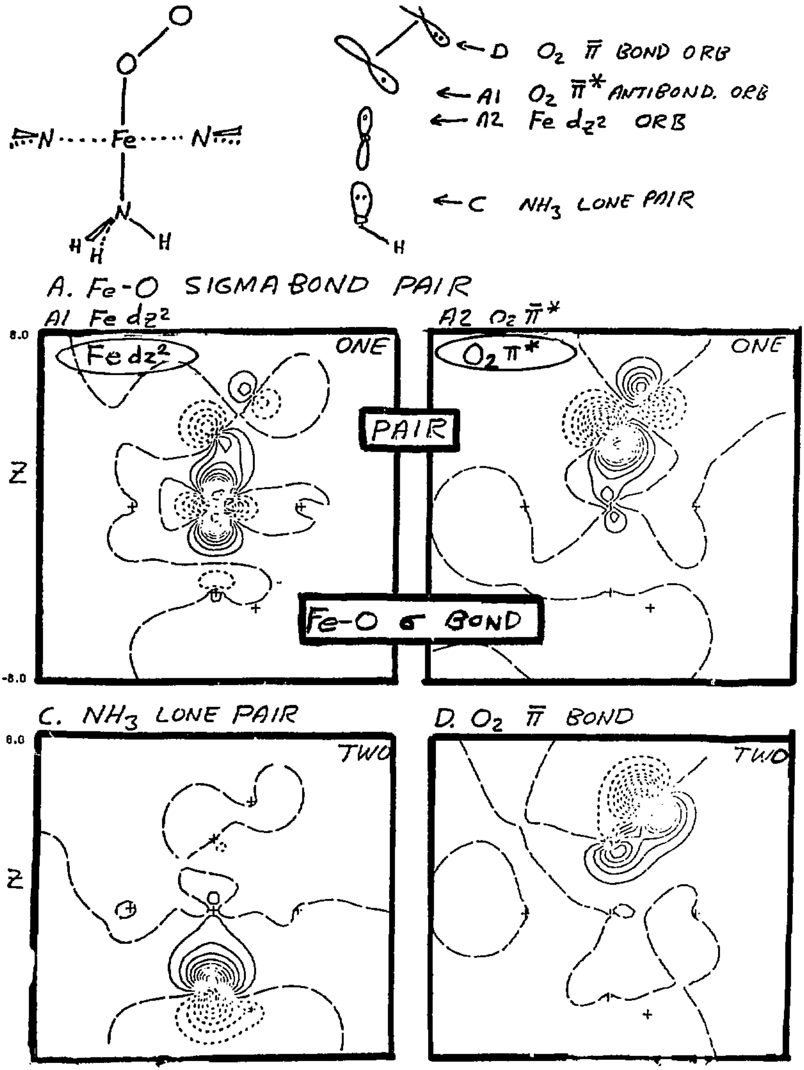
x^2-y^2	↑1 (or ↑)	
z^2	↑1	Lpσ
yz	↑1 ↑↓ (or↑↓)	Lpπ
xz	1L	
xy	1	

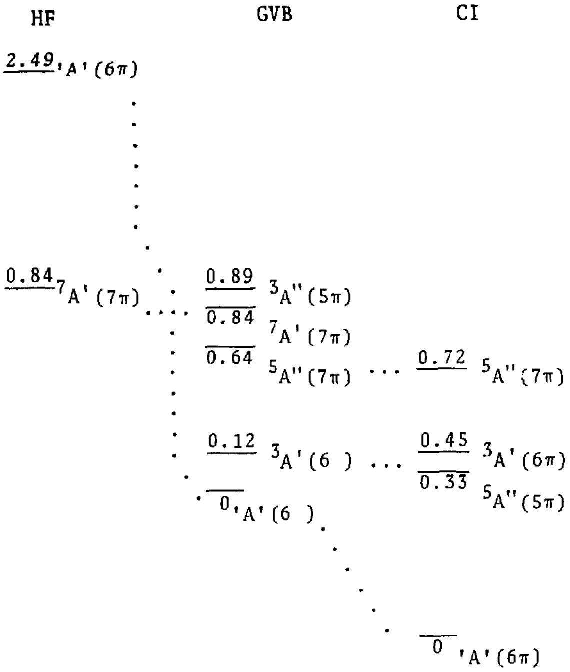
two unpaired spins, therefore paramagnetic

Fe t state O_2

x^2-y^2		
z^2	1	↓ Lpσ
yz	1	↓ Lpπ
xz	1L	
xy	1L	

no unpaired spins, therefore diamagnetic


Fe
N
O
H
D O₂ π BOND ORB
A1 O₂ π* ANTIBOND. ORB
A2 Fe dz² ORB
C NH₃ LONE PAIR
A. Fe-O SIGMA BOND PAIR
A1 Fe dz²
A2 O₂ π*
Fe dz²
O₂ π*
ONE
PAIR
Fe-O σ BOND
8.0
-8.0
Z
C. NH₃ LONE PAIR
D. O₂ π BOND
TWO
6.0



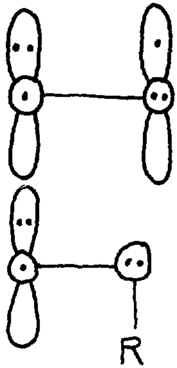
MbO_2 Excitation energy (eV), points

1. HF is bad, bad, bad. Gets septet (S=3) ground state, also triplet and quintet below closed shell singlet.
2. GVB ok, correct ordering.
3. CI needed for accurate E.

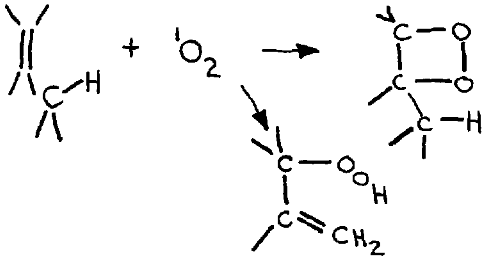
SUMMARY - OXIDATION

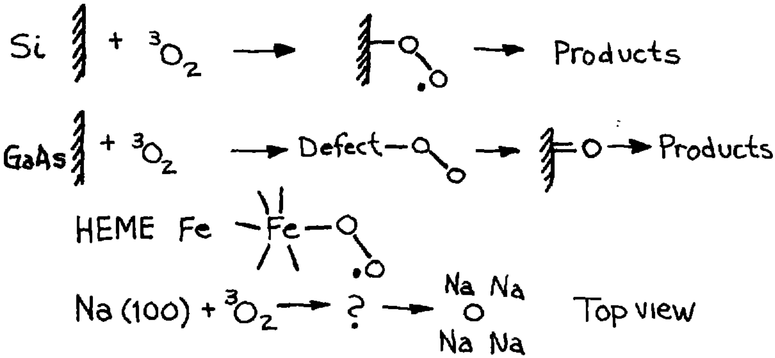
1. $^3O_2$ and $^1O_2$ are biradicals

2. Attack upon radical gets peroxy radical
Radical electron is π, perpendicular to ROO plane



3. Attack $^1O_2$ on olefin





EXAMPLE: Ni_8

Ni_8, simple cubic 80 electrons. Basic configuration energy, d^9 on each Ni = 0. d^8 on each Ni (t_{2g} holes) = 2[illegible].35 eV; (e_g holes) = 23.73 eV. ~d^{10} on each Ni, 32.91 eV, Xα state (start with 44.57 eV, d^{10}, then $e_\mu^2 \rightarrow a_{1g}^2$

Band structure. d^9: d band width, 2.67 eV; s band width, 6.04 eV.

SIMPLIFICATIONS

FACTS

1. When bonded, Ni favors a $(4s)^1(3d)^9$ configuration
2. 3d orbitals are too small and too tightly bound to enter strongly into metal-metal bonds
3. Changing 3d occupation has only a small effect on 4s-like orbitals.

CONCLUSION

Study 4s and 3d bands of solid separately by

1. averaging over all five $(3d)^9$ configurations (to avoid bias)
2. replace this $(3d)^9$ shell with an effective potential.

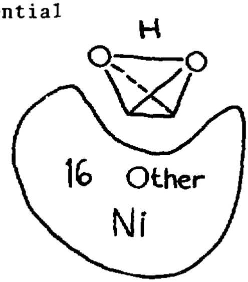
Therefore Ni is reduced from

$$28\ e^-\text{(full atom)} \underset{\substack{\uparrow\\ \text{core effective potential}}}{\rightarrow} 10\ e^-\text{(valence)} \underset{\substack{\uparrow\\ d^9 \text{ averaged potential}}}{\rightarrow} 1\ e^-\ \text{(4s only)}$$

Bonding of H to Bridge Site Of Ni_{20}. Two Cases:

a. Focus on two nearest Ni and two next nearest (tetrahedron of Ni) and use all ten valence electrons; describe other ten Ni using d^9-average potential (therefore 1 elect/Ni)

b. Describe all twenty Ni using d^9-averaged potential



	Case a 4 full Ni 16 1e Ni	Case b 20 1e Ni
Bond energy	66 kcal	63 kcal
Vibrational freq.	1507 cm^{-1}	1428 cm^{-1}
Distance of H above surface	1.03 Å	0.99 Å

Conclusion: good description of chemisorption of H using d^9-averaged potential. This leads to enormous computational savings.

Geometries

S at Ni(100) Ni_5S , 1.36 Å $Ni_{20}S$, 1.24 Å Exper. 1.3±0.1

O at Na(100) Ni_5^+O , 0.96 Å $Ni_{20}O$, 0.88 Å Exper. 0.9±0.1

Vibrational frequency

H at Ni(100) $Ni_{20}H$ 592 cm^{-1} 73 MeV Exper., 605 cm^{-1}

Site	Cluster size	R_{NiH} (Å)	ω_e (cm^{-1})	D_e (kcal)
1 fold	20	1.50	2286	36
2 fold	20	1.59	1428	63
	28	1.55	1419	57
3 fold no atom beneath	28	1.61	1185	52
3 fold atom beneath	20	1.63	1248	74
	28	1.64	1216	72
4 fold	20	1.78	592	70

Experiment: diatomic NiH: R_{NiH} = 1.47 Å

ω_e = 1811 cm^{-1}

Ni(100) (Anderson): ω_e = 605 cm^{-1}

$H_3Ni_4Cp_4$ (Bau): H on 3 fold site

R_{NiH} = 1.691 Å

Geometries

S at Ni(100) —— Ni_5S , 1.36A $Ni_{20}S$, 1.24 A Exper. 1.3 0.1

O at Ni(100) —— Ni_5^+O , 0.96 A $Ni_{20}O$, 0.88 A 0.9 0.1

Vibrational Frequency

H at Ni(100) —— $Ni_{20}H$, 592 cm^{-1} 73 MeV Exper.: 605 cm^{-1}

Bond Energies

S at Ni(100) —— Ni_4S , 3.9 eV $Ni_{20}S$, 4.1 eV Exper." 4.5 eV

therefore, $Ni_{20}S + H_2 \rightarrow Ni_{20} + H_2S$, $\Delta H = +23$ kcal

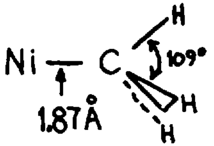
O at Ni(100) —— Ni_5^+O , 3.1 eV $Ni_{20}S$, 2.85 eV

therefore, $Ni_{20}O + H_2 \rightarrow Ni_{20} + H_2O$, $\Delta H = -52$ kcal

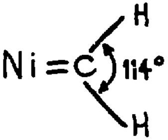
$Ni_{20} + O_2 \rightarrow ONi_{20}O$, $\Delta H = -12$ kcal

CO at Ni(100) —— $Ni_{20}CO$, 27 kcal , Exper.: 30-32 kcal

Ni – C Bonds

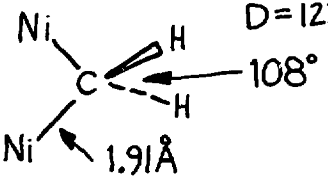


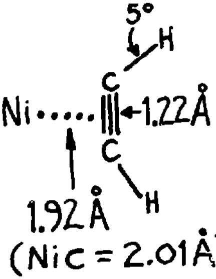
D = 60 KCAL



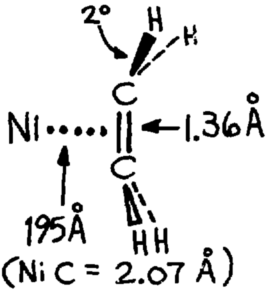
D= 65 KCAL

D=122 KCAL

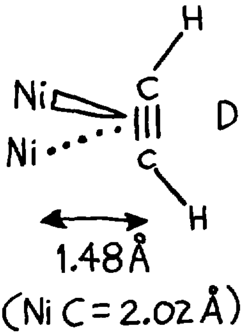




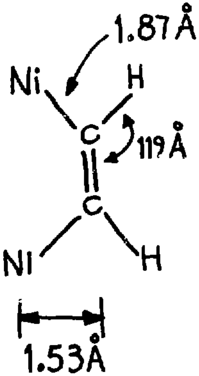
D= 17 KCAL



D=14 KCAL



D=32 KCAL



D=46 KCAL

OPEN SHELL RHF (Restricted Hartree Fock) and MCSCF (Multiconfiguration)

Similar in many concepts

$\{f_i\}$ basis functions

$\downarrow$

$\{\phi_j = \sum x_{ij} f_j\}$ orthonormal molecular orbitals orbitals

$\downarrow$

$$\Psi_{J,\nu} = \mathcal{A}(\phi_{j_1}(1)\bar{\phi}_{j_1}(2)\ldots\bar{\phi}_{j_{k/2}}(k)\,\phi_{\ell_1}(k+1)\ldots\phi_{\ell_{N-k}}(N)\,\chi_\nu$$

J labels space orbital product

$\chi_\nu^{S,M}$ = spin eigenfunction for N-K electrons

$\downarrow$

$$\Psi = \sum_{J,\nu} C_{J,\nu}\, \Psi_{J,\nu}$$

Why RHF?

a. Parent configuration SCF for excited states (relaxation effects)

b. Singlet biradicals; transition states

c. More convenient than UHF for starting CI

WHY MCSCF?

a. Correct dissociation limits

b. Originally hoped to get good D_e

c. Better description of weak bonds

*d. Good orbitals for CI

Notice implied restriction — all configurations are constructed from one mutually orthogonal pool of MO's. RHF only one configuration of this type (or possibly more with $C_{J,\nu}$ fixed by symmetry). The problem: to determine "best" x_{ij} and $C_{J\nu}$ for a given selection of J,ν.

Equivalence Classes of Orbitals

Generally orbitals can be put into subsets with the property that mixing of orbitals within the same subset need not change Ψ. This is true if:

a. Every $\Psi_{J,\nu}$ is unchanged

1. $\{\phi_j\}$ doubly occupied in every $\Psi_{J,\nu}$
2. $\{\phi_j\}$ empty in every $\Psi_{J,\nu}$

b. This set $\Psi_{J,\nu}$ is closed under mixing $\{\phi_j\}$

Partition orbitals into equivalence sets $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3, \ldots$.

Example 1

$$\Psi = (\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\ \phi_3\phi_4\alpha\alpha)$$

(triplet S=1, M=1)

$$\left.\begin{aligned} \phi_1 &\leftarrow (\phi_1 + \varepsilon\phi_2)/\sqrt{1+\varepsilon^2} \\ \phi_2 &\leftarrow (\phi_2 + \varepsilon\phi_1)/\sqrt{1+\varepsilon^2} \end{aligned}\right\} \text{ leaves } \Psi \text{ unchanged}$$

$$\left.\begin{aligned} \phi_3 &\leftarrow (\phi_3 + \varepsilon\phi_4)/\sqrt{1+\varepsilon^2} \\ \phi_4 &\leftarrow (\phi_4 + \varepsilon\phi_3)/\sqrt{1+\varepsilon^2} \end{aligned}\right\} \text{ leaves } \Psi \text{ unchanged}$$

$$\mathcal{D} = \{\phi_1\ \phi_2\},\quad \mathcal{S} = \{\phi_3\ \phi_4\},\quad v = \{\phi_5 \ldots\}$$

E. ample 2

$$\Psi = \mathcal{A}\left(\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\ \phi_3\phi_4\ \frac{\alpha\beta-\beta\alpha}{\sqrt{2}}\right)$$

(singlet S=1, M=0)

$$\phi_3 \leftarrow (\phi_3+\epsilon\phi_4)/\sqrt{1+\epsilon^2}$$

$$\phi_4 \leftarrow (\phi_4+\epsilon\phi_3)\ \sqrt{1+\epsilon^2}$$

$$\Psi_1 \leftarrow \mathcal{A}\left(\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\ \ \phi_3\phi_4\ \frac{\alpha\beta-\beta\alpha}{\sqrt{2}}\right)\left(\frac{1-\epsilon^2}{1+\epsilon^2}\right)$$

$$-\ \sqrt{2}\ \epsilon\ \ (\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\ \ \phi_4\bar{\phi}_4)/(1+\epsilon^2)$$

$$+\ \sqrt{2}\ \epsilon\ \ (\phi_1\phi_1\phi_2\phi_2\ \ \phi_4\phi_4)/(1+\epsilon^2)$$

$$\mathcal{D} = \{\phi_1,\phi_2\}\ ,\ \mathcal{S}_1 = \{\phi_3\}\ ,\ \mathcal{S}_2 = \{\phi_4\}\ ,\qquad = \{\phi_5 \ldots)$$

Notice that for the second example, if

$$\Psi = C_1\ \ \{\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\ \ \phi_4\phi_4(\alpha\beta-\beta\alpha)/\sqrt{2}\,)\}$$

$$+\ C_2\ \ (\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\ \ \phi_3\phi_3)$$

$$+\ C_3\ \ (\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\ \ \phi_4\bar{\phi}_4)$$

then mixing $\phi_3+\phi_4$ brings in no new terms, so

$$\mathcal{D} = \{\phi_1\ \phi_2\}\ ,\ \mathcal{S} = \{\phi_3\phi_4\}\ ,\qquad v = \{\phi_5 \ldots\}$$

<u>but</u> mixing $\phi_3+\phi_4$ for fixed C's would change Ψ.

CONCLUSION: There is no "best" choice of orbitals within each set. E and Ψ will be unchanged (after the best C's are chosen). This freedom can be used as in closed she.. cases to eliminate some Lagrangian multipliers.

The energy expression

$$E = \sum C_{I\nu} C_{J\mu} H_{I\nu,J\mu}$$

assuming all quantities to be real. Variation of $C_{I\mu}$ for fixed orbitals gives

$$\sum_{J\mu} H_{I\mu J\mu} C_{J\mu} = E C_{I\nu}$$

(the matrix eigenvalue problem).

$$H_{I\mu J\nu} = \Gamma_{\mathscr{l}}^{I\mu J\nu} \mathscr{l}$$

where $\mathscr{l}$ is a basic integral over MO's.

$$E = \sum_{\mathscr{l}} \left(\sum C_{I\mu} C_{J\nu} \Gamma_{\mathscr{l}}^{I\mu J\nu} \right) \mathscr{l}$$

$$E = \sum_{\mathscr{l}} Q_{\mathscr{l}} \mathscr{l}$$

Determination of optimal MO's: variation of orbitals for fixed C's

$$\frac{\delta E}{\delta \langle \phi_j |} = \sum Q_j \frac{\delta \mathscr{l}}{\delta \langle \phi_j |}$$

Lagrangian multipliers to ensure orthogonality: minimize

$$E - \sum_{ij} \lambda_{ij} \langle \phi_j | \phi_i \rangle$$

$$\sum Q_{\mathscr{l}} \frac{\delta \mathscr{l}}{\delta \langle \phi_j |} = \sum_i \lambda_{ij} | \phi_i \rangle$$

$$\sum Q_{\mathscr{l}} \frac{\delta \mathscr{l}}{\delta \langle \phi_i |} = \sum_j \lambda_{ij} \langle \phi_j |$$

$\lambda_{ij} = \lambda^*_{ji}$ ensures $(|\phi_i\rangle)^+ = \langle\phi_i|$ and makes these two variations equivalent so that only the first need be considered. These equations are too hard to solve explicitly. The objective is to find a set of equations for improving the orbitals.
This set should

(a) be easily set up and solved, and

(b) rapidly converge when used iteratively.

No scheme has yet been found which is fully satisfactory.
If $\{\alpha_1 \ldots\}$ denotes the set $\{C_1, \ldots x_{12} \ldots\}$ of parameters to be optimized, the Newton-Raphson or Fletcher-Powell-Davidson schemes require $\underset{-}{\delta\alpha} = \underset{\approx}{J}\, \underset{\sim}{\nabla E}$ where $\underset{\approx}{J} \to (\partial^2 E/\partial\alpha_i\partial\alpha_j)^{-1}$. But $\underset{\approx}{J}$ is too <u>large</u> to determine economically, and too <u>non-diagonal</u> to approximate by diagonal elements only, and <u>singular</u> due to non-uniqueness of best α.
The basic approach usually expands n^{th}, iterate in $(n-1)^{th}$

$$(\phi_1^{(n)} \ldots) = (\phi_1^{(n-1)} \ldots)\, U$$

i.e.,

$$\phi_i^{(n)} = \sum \phi_j^{(n-1)} U_{ji}$$

Now suppose we are near convergence, so $U_{ii} \sim 1$ and $U_{ij} \sim$ small.
Consider the simplest mixing process $\phi_i \leftrightarrow \phi_j$

$$\phi_i^{(n)} \sim \phi_i^{(n-1)} + u_{ji}\, \phi_j^{(n-1)}$$

$$\phi_j^{(n)} \sim \phi_j^{(n-1)} - u_{ji}\, \phi_j^{(n-1)}$$

ϕ_i, ϕ_j still orthogonal to first order with fixed coefficients:

$$\psi^{(n)} \cong \psi^{(n-1)} + u_{ji} \sum C_{J,\nu}\, a_j^+ a_i\, \Psi_{J,\nu}$$
$$- \; u_{ji} \sum C_{J,\nu}\, a_i^+ a_j\, \Psi_{J,\nu}$$

$$\delta\psi^{(n)} = u_{ji}(a_j^+ a_i - a_i^+ a_j)\psi^{(n-1)} = u_{ji}\, E_{ji}\, \psi^{(n-1)}$$

See previous example:

$$\psi = \mathcal{A}\{\phi_1\bar{\phi}_1 \ldots \phi_k\phi_{k+1}(\alpha\beta-\beta\alpha)/\sqrt{2}\}$$

$$a_{k+1} a_k\, \psi = \mathcal{A}\{\phi_1 \ldots \phi_{k+1}\phi_{k+1}(\alpha\beta-\beta\alpha)/\sqrt{2}\}$$
$$\equiv \sqrt{2}\, \mathcal{A}\{\phi_1 \ldots \phi_{k+1}\phi_{k+1}\}$$

$$a_k^+ \; a_{k+1}\psi = \sqrt{2}\, \mathcal{A}\{\phi_1 \ldots \phi_k\phi_k\}$$

so,

$$\phi_k \leftarrow \phi_k + \varepsilon\phi_{k+1}$$

$$\phi_{k+1} \leftarrow \phi_{k+1} - \varepsilon\phi_k$$

$$\Downarrow$$

$$\delta\psi = \varepsilon\sqrt{2}\, [\mathcal{A}\{\phi_1 \ldots \phi_{k+1}\phi_{k+1}\} - \mathcal{A}\{\phi_1 \ldots \phi_k\phi_k\}]$$

as found before to first order in ε.

$$E^{(n)} \cong E^{(n-1)} + \langle\delta\psi|H|\psi\rangle + \text{c.c.}$$

$$\frac{\partial E^{(n)}}{\partial u_{ji}} = 2\, \langle(a_j^+ a_i - a_i^+ a_j)\, \psi^{(n)}|H|\psi^{(n)}\rangle$$

Brillouin stationary condition

$$\boxed{\frac{\partial E^{(n)}}{\partial u_{ij}} = 0} \Rightarrow \langle (a_j^+ a_i - a_i^+ a_j)\psi^{(n)}|H|\psi^{(n)}\rangle = 0$$

$$\boxed{\langle E_{ji}\,\psi^{(n)}|H|\psi^{(n)}\rangle = 0}$$

2 × 2 ROTATION METHOD

To find best u_{ij}, assuming all others are zero, we need higher order results:

$$\phi_i^{(n)} = \left(\phi_i^{(n-1)} + u_{ji}\,\phi_j^{(n-1)}\right)\Big/\sqrt{1+u_{ji}^2}$$

$$\phi_j^{(n)} = \left(-u_{ji}\phi_i^{(n-1)} + \phi_j^{(n-1)}\right)\Big/\sqrt{1+u_{ji}^2}$$

$$\phi_i^{(n)} \cong (1 - \tfrac{1}{2}u_{ji}^2)\,\phi_i^{(n-1)} + u_{ji}\,\phi_j^{(n-1)}$$

$$\phi_j^{(n)} \cong (1 - \tfrac{1}{2}u_{ji}^2)\,\phi_j^{(n-1)} - u_{ji}\,\phi_i^{(n-1)}$$

exactly orthogonal, normalized to second order

$$\delta\psi_1^{(n)} = E_{ji}\,\Psi^{(n-1)}$$

$$\delta\psi_2^{(n)} = -\tfrac{1}{2}\,(a_i^+ a_i + a_j^+ a_j)\Psi^{(n-1)} + E_{ji}^2\,\Psi^{(n-1)}$$

$$\delta E = 2u_{ji}\langle E_{ji}\,\Psi^{(n-1)}|H|\Psi^{(n-1)}\rangle$$

$$+ \; u_{ji}^2\left(2\langle\psi^{(n-1)}|H|\delta\psi_2^{(n-1)}\rangle + \langle\delta\psi_1^{(n-1)}|H|\delta\psi_1^{(n-1)}\rangle\right)$$

$$\frac{\partial E}{\partial u_{ji}} = 0$$

$$\Downarrow$$

$$u_{ji} = \frac{<E_{ji}\Psi^{(n-1)}|H|\Psi^{(n-1)}>}{2<^{(n-1)}|H|\delta\psi_2^{(n-1)}>+<\delta\psi_1^{(n-1)}|H|\delta\psi_1^{(n-1)}>}$$

Example of closed (i) $\leftrightarrow$ empty (j) mixing

$$u_{ji} = \frac{<i|\tilde{F}|j>}{<i|\tilde{F}|i> - <j|\tilde{F}|j> + 2J_{12} - 6K_{12}}$$

$$\tilde{F} = n_i(h_{core} + 2\mathcal{J}_i - \mathcal{K}_i) \qquad (n_i=2)$$

or

$$\sqrt{2}\, u_{ji} = \frac{<\Psi_{i\to j}|H|\Psi_o>}{E_o - E_{i\to j} - K_{ij}}$$

This is the correct expression for u_{ij} for 2×2 rotation method. In this method one cycles sequentially through the u_{ij}, mixing one pair of orbitals at a time. Monotonic convergence is assured but cost is high and convergence is slow.

Simultaneous Variations of u_{ij}

The alternative is to vary many of the u_{ij} simultaneously. A true second order method would generalize

$$u_{ij} \simeq -\frac{\partial E/\partial u_{ij}}{\partial^2 E/\partial u_{ij}}$$

to

$$\underset{=}{K}\,\underset{-}{u} = -\begin{pmatrix} \partial E/\partial u_{12} \\ \partial E/\partial u_{13} \\ \vdots \end{pmatrix}$$

where

$$K_{pq} = \frac{\partial^2 E}{\partial u_p\, \partial u_q}$$

This is usually not feasible because $\underset{-}{K}$ is too large to construct or invert.

Approximate K diagonally (ignore coupling between u_{ij})

$$u_{ji} \cong -\frac{\langle E_{ji}\,\Psi^{(n-1)}|H|\psi^{(n-1)}\rangle}{\tilde{K}_{ji,ji}}$$

where

$$\tilde{K}_{ji,ji} \cong \frac{\partial^2 E}{\partial u_{ji}^2}$$

Now from comparison with closed shell SCF derived the same way, we know $K_{ij,ij} \cong \Delta E\, \| E_{ji}\psi \|^2$ so

$$\| E_{ji}\psi \| u_{ji} \cong -\frac{\langle E_{ji}\psi|H|\psi\rangle}{(E_{(E_{ji}\psi)} - E_\psi)\ \| E_{ji}\psi \|}$$

Now the right-hand side is about the same quantity which would be arrived at by a CI calculation, so one can do a "CI" using the set

$$\psi^{(n-1)}\ ,\quad \frac{E_{12}\ \psi^{n-1}}{\| E_{12}\ \psi^{(n-1)} \|}\ \cdots\ \frac{E_{ij}\ \psi^{n-1}}{\| E_{ij}\ \psi^{n-1} \|}\ \cdots$$

which will give (using $C_o=1$)

$$\psi \;=\; \psi^{[n-1]} + \sum c_{ij} \frac{E_{ij}\ \psi^{[n-1]}}{\| E_{ij}\ \psi^{(n-1)} \|}$$

then

$$u_{ji} \;\cong\; \frac{c_{ij}}{\| E_{ij} \psi \|}$$

The CI to some extent properly accounts for K^{-1} coupling between u_{ji}, i.e., it is better (but more costly) to do CI rather than just perturbation theory (Grein & Chang).

Bender has used INO (iterative natural orbital)
This method does CI in the basis

$$\{\ \Psi_I\ ,\quad \Psi_{I,i\to j}\ ,\quad \Psi_{I,j\to i}\ \}$$

which is a much larger CI. Then the NO's of this $\psi^{(n)}$ are used as new orbitals. Ruedenberg has noted that for $\psi^{(n-1)}$ the orbitals in $\psi^{(n-1)}$ may differ from NO's by a transformation

$$\{\chi_1 \chi_2\ \ldots\ \} \;=\; \{\phi_1 \phi_2\ \ldots\ \}\ U$$

while in $\psi^{(n)}$ (CI) they may be

$$\{\chi_1'\ \ldots\} \;=\; \{\phi_1 \phi_2\ \ldots\ \}\ V$$

so an improved set $\{\phi\ \}$ may be given by

$$\{\phi_1' \dots\} = \{\chi_1' \dots\} U^{-1} = \{\phi_1 \dots\} V U^{-1}$$

The closed shell result also suggests that u_{ji} to first order agrees with the eigenvectors of a matrix

$$\tilde{F}_{ij} = \langle E_{ji}\, \psi^{(n-1)} | H | \psi^{(n-1)} \rangle \qquad i \neq j$$

$$\tilde{F}_{ii} = \varepsilon_i \, \| E_{ji} \psi^{(n-1)} \|^2$$

provided one can define diagonal elements so that

$$\frac{\partial^2}{\partial u_{ij}^2} = \tilde{F}_{ii} - \tilde{F}_{jj}$$

Note for K basis functions, there are $K(K-1)/2$ u_{ij}, but only K $\tilde{F}_{ii}$, so it is not clear that a matrix $\tilde{F}$ can be formed whose eigenvectors give a good approximation to u_{ij}. If some $\tilde{F}_{ii}$-$\tilde{F}_{jj}$ have wrong sign, oscillation would be expected.

Fock Operator Method

This equation involving the Fock operator is usually derived somewhat differently. Consider an orthonormal basis $\{|f\rangle\}$ and the MO basis $\{|\phi_1\rangle \dots\} = \{|f\ \rangle \dots\} \underset{\sim}{U}$, i.e.,

$$|\phi_i\rangle = \sum_j U_{ji} \, |f_j\rangle$$

Then one can consider variations in u_{ji} subject to orthogonality:

Minimize

$$E - \sum \lambda_{ji} \langle\phi_i|\phi_j\rangle$$

$$\frac{\delta E}{\delta\langle\phi_i|} = \sum_j F_{ij}|\phi_j\rangle = \sum \lambda_{ji} |\phi_j\rangle$$

$$\lambda_{ij} = \lambda_{ji}^*$$

Detail:

$$h_{ij} = \langle\phi_i|h|\phi_j\rangle$$

$$h = K.E. + V_{Ne}$$

$$[ij\|k\ell] = \int \frac{\phi_i^*(1)\phi_j(1)\ \phi_k^*(2)\phi_\ell(2)}{r_{12}} d\tau_1 d\tau_2$$

$$= \langle\phi_i\phi_k| \frac{1}{r_{12}} |\phi_j\phi_\ell\rangle$$

$$\sum Q_{\ell}\mathcal{l} = \sum_{ij} \rho_{ij}h_{ji} + \sum_{ii'jj'} \Gamma_{ii'jj'}[ji\|j'i']$$

where ρ_{ji} = density matrix in MO basis

$$\frac{\delta \sum Q_{\ell}\mathcal{l}}{\delta\langle\phi_j|} = \sum_i \rho_{ij}\ h|\phi_i\rangle + \sum_{ii'j'} \Gamma_{ii'jj'} \frac{\delta[ji\|j'i']}{\delta\langle\phi_j|}$$

$$+ \sum_{ii'j'} \Gamma_{ii'jj'} \frac{\delta[j'i\|ji']}{\delta\langle\phi_j|} = \sum F_{ji}|\phi\ \rangle$$

$$\boxed{\sum_i F_{ji}|\phi_i\rangle = \sum_i \lambda_{ij}|\phi_i\rangle}$$

Careful consideration of the Lagrangian multipliers shows

$$\lambda_{ij} = \lambda^*_{ji}$$

is necessary to ensure $(|\phi_i\rangle)^\dagger = \langle\phi_i|$, if $|\phi_i\rangle$ and $\langle\phi_i|$ are varied independently. Hence the condition for a stationary E is

$$\left\{ \begin{array}{c} \sum_k F_{jk}|\phi_k\rangle = \sum_i \lambda_{ij}|\phi_i\rangle \\ \\ \lambda_{ij} = \lambda^*_{ji} \end{array} \right\}$$

($F_{ij} = F^\dagger_{ji}$ follows from definition.)

"Canonical" Orbitals

$$\lambda_{ij} = 0 \quad \text{if} \quad i,j \text{ in same set } \mathscr{W} .$$

Evaluation of λ_{ij}:

$$\lambda_{ij} = \sum_k \langle\phi_i|F_{jk}|\phi_k\rangle = \langle\phi_i|F_j|\phi_j\rangle$$

$$\lambda^*_{ji} = \sum_k \langle\phi_k|F_{ki}|\phi_j\rangle = \langle\phi_i|F_i|\phi_j\rangle$$

$$\lambda_{ij}-\lambda^*_{ji} = 0 = \langle\phi_i|F_j-F_i|\phi_j\rangle$$

$$\boxed{\langle\phi_i|F_j-F_i|\phi_j\rangle \equiv \langle E_{ji}\psi|H|\psi\rangle}$$

$$\left(F_j|\phi_j\rangle - \sum_{i\neq j} \lambda_{ij}|\phi_i\rangle\right) = \lambda_{jj}|\phi_j\rangle$$

$$\lambda_{ij} = \langle\phi_i|F_j|\phi_j\rangle\,(1-b_{ij}) + \langle\phi_i|F_i|\phi_j\rangle\, b_{ij}$$

$$G_{kj} = \langle\phi_k|F_j - F_k|\phi_j\rangle\, b_{kj} = \langle\phi_k|F_j|\phi_j\rangle - \lambda_{kj}$$

$$G_{jj} = \langle\phi_j|F_j|\phi_j\rangle$$

<u>Simple Example</u> – one open shell RHF doublet state.

$$F_{\mathcal{D}} = 2\left(h + 2\mathcal{J}_{\mathcal{D}} - \mathcal{K}_{\mathcal{D}} + \mathcal{J}_{\mathcal{S}} - \tfrac{1}{2}\mathcal{K}_{\mathcal{S}}\right)$$

$$F_{\mathcal{S}} = h + 2\mathcal{J}_{\mathcal{D}} - \mathcal{K}_{\mathcal{D}} + \xi\left(\mathcal{J}_{\mathcal{S}} - \mathcal{K}_{\mathcal{S}}\right)$$

(ξ arbitrary)

Define

$$\left.\begin{aligned} F_\alpha &= h + 2\mathcal{J}_{\mathcal{D}} - \mathcal{K}_{\mathcal{D}} + \mathcal{J}_{\mathcal{S}} - \mathcal{K}_{\mathcal{S}} \\ F_\beta &= h + 2\mathcal{J}_{\mathcal{D}} - \mathcal{K}_{\mathcal{D}} + \mathcal{J}_{\mathcal{S}} \end{aligned}\right\} \text{ UHF operators}$$

$$F_{\mathcal{D}} = F_\alpha + F_\beta \quad , \qquad F_\alpha|\mathcal{S}\rangle = F_{\mathcal{S}}|\mathcal{S}\rangle$$

$$\tfrac{1}{2}\,\frac{\partial E}{\partial(\mathcal{D}v)} = \langle E_{\mathcal{D}v}\psi|H|\psi\rangle = \langle v|F_{\mathcal{D}}|\mathcal{D}\rangle$$

$$\tfrac{1}{2}\,\frac{\partial E}{\partial(\mathcal{S}v)} = \langle E_{\mathcal{S}v}\psi|H|\psi\rangle = \langle v|F_\alpha|\mathcal{S}\rangle$$

$$\tfrac{1}{2}\,\frac{\partial E}{\partial(\mathcal{D}\mathcal{S})} = \langle E_{\mathcal{D}\mathcal{S}}\,\psi|H|\psi\rangle = \langle\mathcal{S}|F_\beta|\mathcal{D}\rangle$$

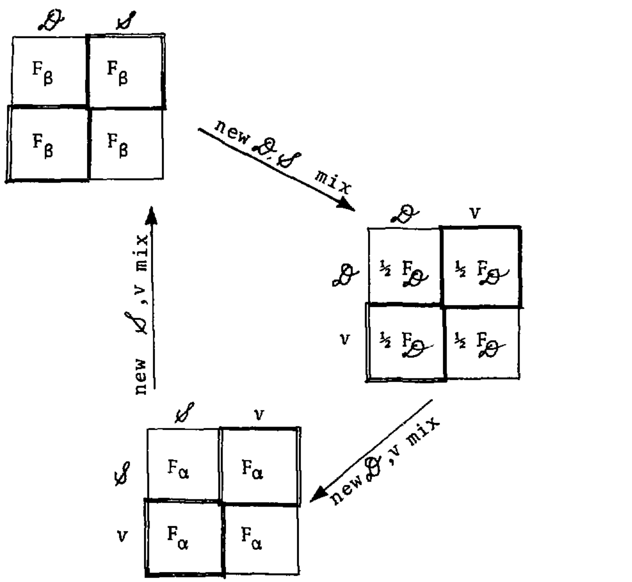
$$\tfrac{1}{2}\,\frac{\partial^2 E}{\partial(\mathcal{D}v)^2} = \langle v|F_{\mathcal{D}}|v\rangle - \langle\mathcal{D}|F_{\mathcal{D}}|\mathcal{D}\rangle + 6K_{\mathcal{D}v} - 2J_{\mathcal{D}v}$$

$$\tfrac{1}{2}\,\frac{\partial^2 E}{\partial(\mathcal{S}v)^2} = \langle v|F_\alpha|v\rangle - \langle\mathcal{S}|F_\alpha|\mathcal{S}\rangle + K_{\mathcal{S}v} - J_{\mathcal{S}v}$$

$$\tfrac{1}{2}\,\frac{\partial^2 E}{\partial[\mathcal{D}\mathcal{S}]^2} = \langle\mathcal{S}|F_\beta|\mathcal{S}\rangle - \langle\mathcal{D}|F_\beta|\mathcal{D}\rangle + K_{\mathcal{S}\mathcal{D}} - J_{\mathcal{S}\mathcal{D}}$$

<u>BINKLEY, POPLE, DOBOSH METHOD</u>

(3 Hamiltonian, 2 × 2 rotation)



Zero at convergence (Brillouin). Other blocks chosen to mimic K^{-1} optimum mixing?

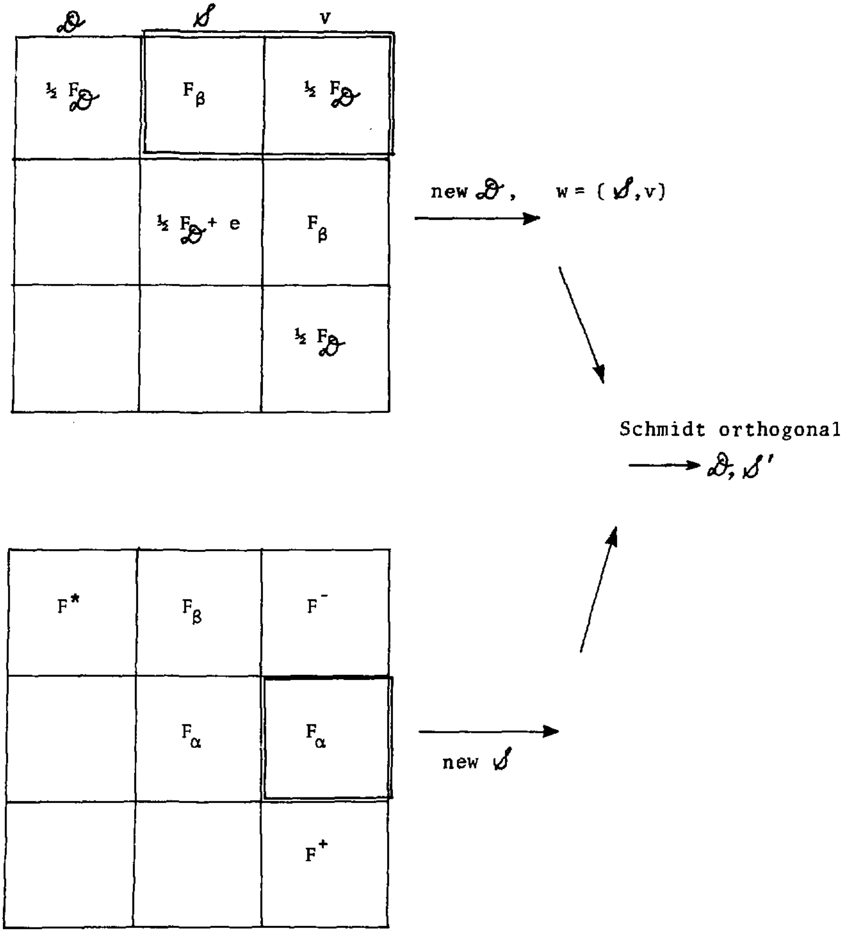
One Hamilton Method — Guest and Saunders

	$\mathcal{D}$	$\mathcal{S}$	v
$\mathcal{D}$	F_1	F_β	$\frac{1}{2} F_{\mathcal{D}}$
$\mathcal{S}$	F_β	F_2	F_α
v	$\frac{1}{2} F_{\mathcal{D}}$	F_α	F_3

- Zero at convergence (Brillouin)
- F_1, F_2, F_3 arbitrary
- $F_3 = F_\alpha$, $F_1 = F_\beta + c\mathbf{1} = F_2$

common choice

Roothaan-Bagus — 2 Hamiltonian Method



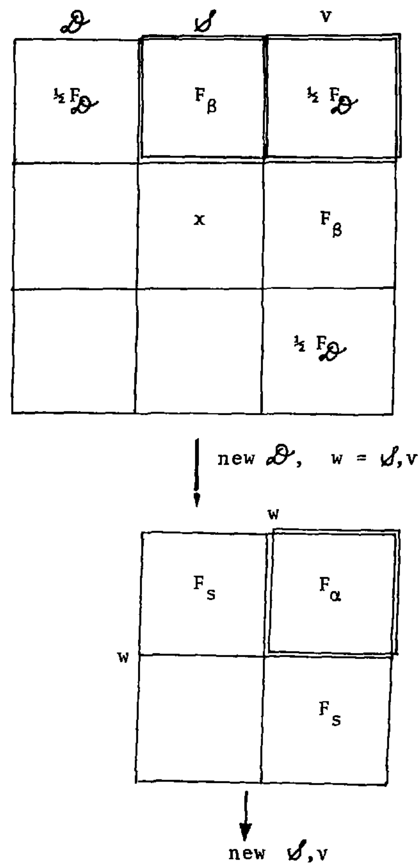
- Disadvantage:

  Schmidt orthogonalization.
  Strange choices in arbitrary blocks.

- Advantage:

  One integral read/cycle.

Davidson (improved) — 2 Hamiltonian method.



- Disadvantage: still some strange choices in diagonal.
- Advantage: one integral read/cycle.

NATURAL ORBITALS

$$\rho = N\int \psi(1,2,\ldots N)\ \psi^*(1',2,\ldots N)d\tau_2\ldots d\tau_N$$

$$= \rho(1,1') = \sum \rho_{ij}\ \phi_i(1)\phi_j^*(1')$$

$$\int \rho(1,1')\chi(1')d\tau_1' = \mu\chi(1) \quad \text{for NO}$$

$$\rho(1,1') = \sum \mu_i\ \chi_i(1)\chi_i^*(1')$$

ρ is now "diagonal".

ψ generally more rapidly convergent; μ_i = occupation number,

$$\mu_i = \sum_{i\in K} |C_K|^2\ M_{i,K}$$

2e$^-$ Wave Function Special Case

$$\psi = \sum_{j,i=1}^{K} C_{ij}\ \phi_i(1)\phi_j(2)$$

$$\downarrow$$

$$\psi = \sum_{i=1}^{K} \mu_i^{\frac{1}{2}}\ \chi_i(1)\chi_i^*(2) \qquad \text{is diagonal}$$

$$\psi_{VB} = C_1(f_Af_B + f_Bf_A) + C_2(f_Af_A) + C_2(f_Af_A)$$

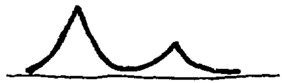
$$\downarrow$$

$$\psi_{VB} = \tilde{C}_1(\chi_1(1)\chi_1(2)) + \tilde{C}_2(\chi_2(1)\chi_2(2))$$

$$\int \chi_1 \chi_2 \, d\tau = 0$$

Bond

$$\chi_1 = a\,f_A + b\,f_B$$



Anti-bond

$$\chi_2 = -b' f_A + a' f_B$$

<u>GVB</u>

$$\psi \cong \mathcal{A}\{g_1(1,2)\,g_2(3,4)\,g_3(5,6)\ldots\}$$

$$g_i = \begin{cases} \chi_i(1)\bar{\chi}_i(2) & \text{core} \\ a\chi_i(1)\bar{\chi}_i(2) - b\chi_i'(1)\bar{\chi}_i'(2) & \text{valence} \end{cases}$$

$\{\chi_i\}$ mutually orthonormal set.

Special MC-SCF form which takes advantage of NO form of $2e^-$ wavefunction. For this assumed form of ψ, ρ is diagonal, i.e.,

$$\psi \rightarrow \rho = \sum \mu_i \, \chi_i \chi_i$$

1. INTRODUCTION

The goal of any semi-empirical or approximate method is to strive for some compromise between ease of application and accuracy. Ease of application generally implies dropping terms difficult to evaluate. Within any given level of a theory parameters can be introduced to mimic the behavior of that theory applied from first principles (*ab initio*). These parameters can also be introduced from experimental observables, or to reproduce experimental observables, or both. In the first case, parameterization can yield results no better than the *ab initio* theory, by design. In the second case, when recourse is made to experiment for some of the parameters, the model constructed, if carefully designed, may shadow a more exacting theory.

In a very real sense successful quantum mechanical models for calculating electronic structure that rely on semi-empirical parameters based on experimental observables contain information about electron correlation — as well as implying a perfect basis set and an "exact" theory. If the model is at the restricted Hartree Fock level, then this information is entirely implicit; if it is a model beyond Hartree-Fock, then some of the correlation effects will be explicit, some implicit. A more successful model, however, will have carefully separated (perhaps unknowingly) the explicit from the implicit and as such its allusion to a higher order theory should be derivable from the lower order *ab initio* theory from which it was designed.

Within these notes I would like to discuss semi-empirical quantum mechanical models which explicitly contain correlation corrections. Such models have a twofold utility. Although principally designed to explain experiment, they are useful in examining the correlation problem itself. The quantum mechanics of two or more electrons can be examined by but two theories (themselves approximate); variational theory and perturbation theory. For these notes I have chosen one example of the application of each. The first of these is the very effective Zero Differential Overlap (ZDO) theories used in molecular spectroscopy. Here the variational principle is applied to an approximate Hamiltonian to yield the energy differences between ground and excited states. The second of these methods develops the ground state energy of a system from a reference of doubly occupied bonding orbitals principally through perturbation theory. This theory is rapid and apparently accurate enough to allow the calculation of geometric conformation of very large systems.

2. MOLECULAR ORBITAL THEORY AND BEYOND

The review that follows is by necessity brief, but should serve useful in defining the nomenclature and contrasting the philosophy of the methods discussed.

Molecular electronic structure theory generally begins with molecular orbital theory summarized in the equations below:

Molecular Electronic Schrodinger Equation

$$H\psi_\alpha(1,2,...n) = E_\alpha\psi_\alpha(1,2,...n) \tag{1}$$

Molecular Orbital (MO) Approximation

$$\psi_\alpha \approx \psi_\alpha^o(1,2,...n) = \theta_S \mathcal{A}[\phi_1(1)\phi_2(2)...\phi_n(n)] \tag{2}$$

Linear Combination of Atomic Orbital Approximation (LCAO)

$$\phi_i(j) = \sum_{\mu=1}^{N} \chi_\mu(j)\, C_{\mu i} \tag{3}$$

Variational Principle

$$\frac{\langle \psi_\alpha^o | H | \psi_\alpha^o \rangle}{\langle \psi_\alpha^o | \psi_\alpha^o \rangle} \geq E_x \tag{4}$$

Secular Equation

$$(\underline{F} - \varepsilon_i \underline{\Delta})\underline{C}_i = 0 \tag{5}$$

Self-Consistent Field (SCF)

$$\underline{C}^o \to \underline{F}(\underline{C}^o) \to \underline{C}^1 \to \tilde{\underline{C}}^1 \to \underline{F}(\tilde{\underline{C}}^1) \to \underline{C}^2 \to \ldots \tag{6}$$

The molecular orbital approximation itself, given in Eq. (2), sets the trial wave function equal to a spin projected (θ_S) antisymmetrized product ($\mathcal{A}$) of orbitals $\phi_i(j)$. The assignment of electrons to these orbitals, perhaps by an aufbau principle, creates the reference configuration. The accuracy and consistency of Eq. (2) creates the correlation "problem".

Application of the variational principle with the trial wave function of Eqs. (3) and (4) yields the secular Hartree-Fock Roothaan[1] equation (5), with $\underline{C}_i$ the MO coefficients of ϕ_i, Eq. (3), ε_i the molecular orbital energy, $\underline{\Delta}$, the orbital overlap matrix

$$\Delta_{\mu\nu} = (\chi_\mu|\chi_\nu) = (\mu|\nu) \tag{7}$$

and $\underline{F}$, the Fock or energy matrix, with elements for a closed shell system given by

$$F_{\mu\nu} = (\mu|h|\nu) + \sum P_{\sigma\lambda}[(\sigma\lambda|\mu\nu) - \tfrac{1}{2}(\sigma\mu|\lambda\nu)]$$

$$h = -\frac{\nabla^2}{2} - \sum_A \frac{Z_A}{R_A}$$

$$P_{\sigma\lambda} = \sum_a C_{\sigma a}\, C_{\lambda a}\, n_a$$

$$(\sigma\lambda|\mu\nu) = \langle\sigma\mu|\lambda\nu\rangle = \int d\tau_1 d\tau_2\ \chi^*_\sigma(1)\,\chi^*_\mu(2)\ \gamma^{-1}_{12}\ \chi_\lambda(1)\,\chi_\nu(2)$$

The Fock matrix represents the kinetic energy of the electrons, the nuclear electron attraction, and the repulsion of an electron by the average field of all the electrons. This latter manifests itself through the dependence of Eq. (8) on $\underline{P}$, the first order density matrix (or the charge and bond order matrix if $\underline{\Delta}=\underline{1}$), which in turn depends on $\underline{C}$, the MO coefficients not known and n_a, the orbital occupancy of MO, ϕ_a. This suggests the iterative SCF procedure of Eq. (6) where $\tilde{\underline{C}}^n$ is equal to $\underline{C}^n$ or is extrapolated from it to hasten convergence.

Equations (1) to (6) define the Hartree-Fock LCAO-MO-SCF method. In spite of the many approximations made it is the most successful theory in Quantum Chemistry.

From the computational standpoint, the evaluation of Eqs. (8) requires N^4 difficult two-electron integrals, where N is the number of atomic orbitals in Eq. (3). Although there is hope that this N^4 dependence would decrease to N^3 as systems grow larger by neglecting small integrals, this reduction in N dependence has not yet been realized in any method that can still truly be called *ab initio*.[2,3] In fact, the inclusion of symmetry in molecular programs to simplify the execution of molecular orbital theory often prevents dropping any integrals as a consequence of positioning at least one member of a symmetry-adapted orbital near a member of another. In addition to this N^4 dependence in the evaluation of the integrals there is the repeated N^3 problem associated with the solutions of Eq. (5). As Eq. (5) is the spiritual bottleneck of any molecular orbital method, all semi-empirical methods strive to reduce the number of integrals evaluated to N^3 or less. In the ZDO methods to be discussed in these notes the number of integrals to be evaluated is proportional to N^2.

From the theoretical point of view, the Hartree-Fock SCF-LCAO-MO theory briefly described has one important flaw: for systems of two or more electrons it is incapable of yielding correct results! The reason is that the restricted Hartree-Fock theory calculates the repulsion of an electron in the average field of all the others, and does not correlate the individual trajectories of electrons. This problem is well understood as is the method to correct this shortcoming. A new trial wavefunction is created that is a linear combination of configurations

$$\psi^T = d_o\psi^o + \sum d_{ia}\,\psi_i^a + \sum d_{ijab}\,\psi_{ij}^{ab} + \ldots \tag{9}$$

in which ψ_i^a, ψ_{ij}^{ab}, etc., represent configurations in which one or more electrons have been excited from orbitals ϕ_i, ϕ_j, etc., occupied in the reference configuration ψ_o, into orbitals $\phi_a, \phi_b, \ldots$ orthogonal to the occupied set. Invoking the variational principle yields

$$(\underline{H} - E_\alpha \underline{1})\underline{D}_\alpha = 0 \tag{10}$$

where $\underline{D}_\alpha$ are the expansion coefficients of Eq. (9), $H_{ij} = \langle\psi_{ij..}^{ab..}|H|\psi_{k\ell..}^{cd..}\rangle$ and E_α is a bound to the energy of ψ_α. Equation (10) is the configuration interaction problem. The expansion of Eq. (9) is slow to converge, and most of modern theoretical quantum chemistry (as opposed to applied) is concerned in solving or approximating Eq. (10). Nevertheless, Eq. (10) is capable, as far as we understand, of yielding results that can systematically be improved to approach experiment.

3. SPECTROSCOPIC ZERO DIFFERENTIAL OVERLAP THEORIES

A. Theory

A great deal of the development of quantum mechanics itself can be associated with the observed line spectra of atoms. In an analogous fashion, most modern quantum chemical methods can be traced to attempts to explain molecular spectra. One of the first successful methods in organizing molecular spectra was the semi-empirical method of Pariser, Parr and Pople (PPP) applied to the pi electron system of conjugated hydrocarbons.[4] In addition to yielding results in near quantitative agreement with experiment, the model predicted very directly the importance of configuration interaction at a time when not even the Hartree-Fock molecular orbital theory was well understood. The extension of PPP theory to almost planar systems, to systems in which heteroatoms or substituents polarized the sigma electrons not explicitly considered, and to systems containing transition metals, proved difficult. Although there are several versions designed to extend the domain of applicability,[5] the PPP theory is perhaps best used today to examine new approaches to the correlation problem itself, by providing a well defined model Hamiltonian that is easy to evaluate and to examine.

Semi-empirical all-valence electron methods have been introduced that now execute on a computer at the Hartree-Fock level nearly as rapidly as the PPP method. These methods are applicable to planar or non-planar systems, take into account sigma electron rearrangements in a natural fashion, and have been extended to transition elements, thus eliminating the three major shortcomings of the PPP theory.

Of the many all-valence electron theories that have been proposed, only those of the ZDO type have been systematically applied to the study of molecular spectra. Simpler theories that do not refer directly to the two-electron nature of the Fock matrix, Eq. (7), are difficult to apply to detailed spectroscopy, as spectroscopy is inherently a two-electron phenomenon. On the other hand, theories that purport to more accurate Hamiltonians, and thus more integrals, eventually strike the N^5 (or N^6) integral transformation bottleneck, making the value of any approximations questionable.

In 1965 Pople and co-workers introduced their Complete Neglect of Zero Differential Overlap Method,[6] summarized by the following equations:

A) Rotational Invariance

$$\gamma_{\mu\nu} = \gamma_{AB} = (S_A S_A | S_B S_B) \qquad (11A)$$

B) Core Integrals

$$U_{\mu\mu}^{AA} = \left(\mu \middle| -\frac{\nabla^2}{2} - \frac{Z}{R} \middle| \mu\right) = -I_A - (Z_A^{\nu} - 1)\gamma_{AA} \qquad \text{(CNDO/1)}$$

or (11*B*)

$$= -\frac{(I_A + A_A)}{2} - (Z_A^{\nu} - \tfrac{1}{2})\gamma_{AA} \qquad \text{(CNDO/2)}$$

C) ZDO: $\langle\mu\nu| = \delta_{\mu\nu}\langle\mu\mu|$ (11*C*)

D) $$\nu_{AB} = (S_A|Z_B^\nu/R_B|S_A) = Z_B^\nu \gamma_{AB} \quad (11D)$$

E) $$H_{\mu\nu}^{AB} = \frac{\Delta_{\mu\nu}(\beta_A + \beta_B)}{2} \quad (11E)$$

Approximation C), the ZDO approximation, reduces the number of integrals from N^4 to N^2, yielding for $F_{\mu\mu}$, for example,

$$F_{\mu\mu}^{AA} = U_{\mu\mu}^{AA} - \sum_{B\neq A}(Z_B - P_{BB})\gamma_{AB} + (P_{AA} - \tfrac{1}{2}P_{\mu\mu})\gamma_{AA} \ .$$

The relation between core integral, $U_{\mu\mu}$, ionization potential I_μ, and electron affinity A_μ is derived through the ZDO expression for the valence electron energy of the atomic configuration $s^\ell p^m$,

$$E(s^\ell p^m) = \ell U_{ss} + mU_{pp} + \frac{Z_A^\nu(Z_A^\nu - 1)\gamma_{AA}}{2} \quad (12)$$

where Z_A^ν, the core charge, is equal to $\ell+m$ in the neutral atom, or the number of valence electrons explicitly considered in the calculation.[7] The original methods were parameterized through β_A to give agreement with model minimum basis set *ab initio* work.

In many ways the CNDO method is a natural extension of PPP theory to sigma system, and, as such, was quickly adopted to spectroscopic purposes using the experience gained with pi electron theories.[8] The first of the required modifications was the introduction of the Pariser approximation,[9] or scaling down of the two-electron integrals. Using Eq. (12) one derives

$$\gamma_{\mu\mu}(s) = I_\mu - A_\mu \qquad ,$$

and one may define

$$\gamma_{AA}(s) = \frac{\left(\gamma_{ss}(s) + \gamma_{pp}(s)\right)}{2} \quad . \tag{13}$$

In general,

$$\gamma_{AA}(s) \approx \gamma_{ss}(s) \approx \gamma_{pp}(s) \approx 0.7\gamma_{ss} \qquad \text{(Theoretical)}$$

In order to smoothly connect $\gamma_{AA}(s)$ with R_{AB}^{-1}, the long-range behavior of the two-electron Coulomb integral $(S_A S_A | S_B S_B)$ one generally assumes

$$\gamma_{AB}(s) = \frac{1}{[R_{AB}^{\,n} + \rho_A^{\,n}]^{1/n}} \tag{14}$$

$$\rho_A = \frac{2}{\gamma_{AA}(s) + \gamma_{BB}(s)}$$

For n=1, the Mataga Nishimoto[10] formula is obtained; for n=2, the Ohno-Klopman formula.[11,12]

A second important modification for spectroscopic studies is introduced into Eqs. (11) to correct for the improper placement of the pi molecular orbitals within the sigma. Although the original purposes to which the CNDO model were put seemed relatively insensitive to reasonable error between orbital energies, the calculation of spectra is crucially dependent on these energies. In Eq. (*11E*), the orbital overlap $\Delta_{\mu\nu}$ is replaced by a scaled overlap, $\overline{\Delta}_{\mu\nu}$

$$\begin{aligned} \overline{\Delta}_{ss'} &= \Delta_{ss'} \\ \overline{\Delta}_{sp'} &= \Delta_{s\sigma'}\, g_{s\sigma'} \\ \overline{\Delta}_{pp'} &= \Delta_{\sigma\sigma'}\, f_{\sigma\sigma}\, g_{\sigma\sigma'} + \Delta_{\pi\pi'}\, f_{\pi\pi}\, g_{\pi\pi'} \end{aligned} \tag{15}$$

where $g_{\mu\nu}$ are the Eulerian transformations used to transform from the molecular coordinate system to the local diatomic system (and are required for the calculation of all integrals over Slater-type orbitals) and,[8,13,14]

$$\begin{aligned} f_{\pi\pi} &\approx 0.60 \\ f_{\sigma\sigma} &\approx 1.0 \text{ or } 1.3 \end{aligned} \tag{16}$$

It should be pointed out that Eq. (11*E*) has no theoretical justification in the context of ZDO calculations, and is the weakest point of the theory, as the method creates a strong dependence on a given atomic orbital basis set. The introduction of the two parameters of Eq. (16) free the evaluation of $H_{\mu\nu}$ somewhat from any choice of basis. Although the dependence of $H_{\mu\nu}$ on $R_{\mu\nu}$ is still principally that of $\Delta_{\mu\nu}$, "non-nearest neighbor" $H_{\mu\nu}$ are generally very small.

The Intermediate Neglect of Differential Overlap Theory (INDO) differs from the CNDO theory by refinement of one-center two-electron integrals,[15] viz.,

$$
\begin{aligned}
(SS|SS) &= F^{o} \\
(XX|XX) &= F^{o} + \frac{4F^{2}}{25} \\
(XX|YY) &= F^{o} + \frac{2F^{2}}{25} \qquad (17) \\
(XY|XY) &= \frac{3F^{2}}{25} \\
(SX|SX) &= \frac{G^{1}}{3}
\end{aligned}
$$

In the CNDO theory the higher Slater Gordon factors[16] F^{K} and G^{K} for $K > 0$ are set to zero. Equations (17) include one-center exchange terms and would seem crucial for inclusion in any spectroscopic theory for it is these integrals that split the atomic term energies for a given configuration, and this splitting is large. In practice, the improvement for molecules containing hydrogen and first-row elements is not large, and mostly confined to $(n - \pi^{*})$ excitations. The reason for the improvement has something to do with the inclusion of the integrals that split the $^{1}(n,\pi)$ from $^{3}(n,\pi)$ excitation energies that are set to zero under the CNDO approximations, but probably has more to do with the $f_{\sigma\sigma} = 1.3$ in the INDO spectroscopic programs as presently implemented[13] versus $f_{\sigma\sigma} = 1.0$ in the CNDO versions,[8] Eqs. (15) and (16). The reason that the improvement is not large is the non-atomic-like spin and angular momentum coupling present in the atoms of the first row when they form molecules.

Unlike atoms of the first row, transition metals display a great deal of atomic character in complexes and the INDO/S theory shows a clear superiority over the CNDO/S theory. In these cases though, a clear definition of what is meant by INDO must be given, as many more integrals occur between s, p, and d orbitals than the five given in Eqs. (17). Several investigators[17,18] gave defined INDO to mean only those integrals of the form $(ij|ij)$, i.e., exchange or Coulomb type, in an obvious but simplified extension of Eqs. (17). Although ignoring the general integral $(ij|k\ell)$ introduces rotational variances, these effects appear small. Other investigators have suggested averaging over classes of $(ij|ij)$ integrals to remove these rotational variances,[19] but these definitions will not suffice if one wants a method for spectroscopy. The ignored terms are essential for an accurate estimate of the correlation energy, and for the splitting of various excited states. An example is given for ferrocene in Figure 1.[20]

Both the CNDO/S and the INDO/S model are parameterized on spectroscopic results obtained through Hartree-Fock theory plus extensive "singles only" CI. The energies of the states are usually taken as the roots of Eq. (10). Occasionally an

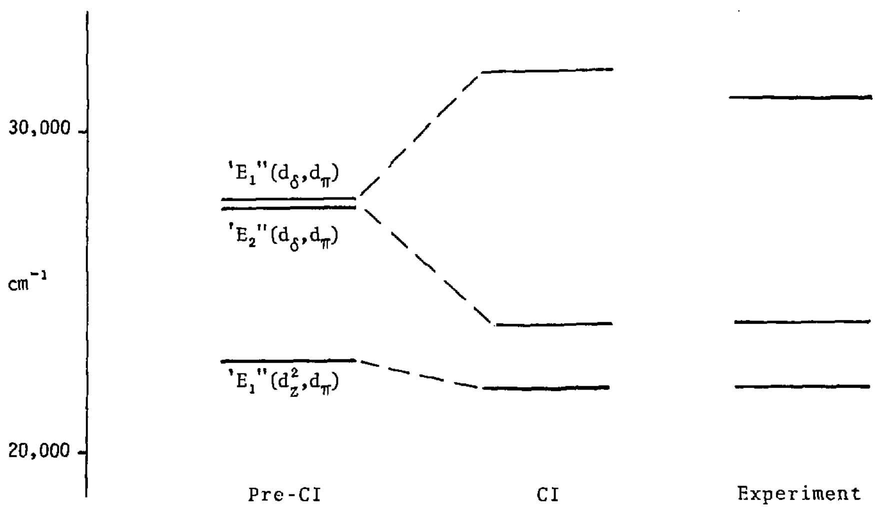


Fig. 1. The (d,d) transitions of ferrocene. The large CI splitting between $^1E_1''$ and $^1E_2''$ is caused by integrals that are not of the classical Coulomb or exchange type.

extrapolation procedure between ΔE(SCF) and CI has been applied to improve results.[21] In both the CNDO/S and INDO/S methods oscillator strengths are estimated using the dipole length operator, including the one-center "charge" terms and the one-center polarization terms, i.e.,

$$\langle\psi_o|\vec{\mu}|\psi_i^a\rangle = \sqrt{2}\,\langle\phi_i|\vec{\mu}|\phi_a\rangle = \sqrt{2}\sum_A\sum_{\alpha}^{A} C_{\alpha i}C_{\alpha a}\vec{R}_A + \sqrt{2}\sum_A\sum_{\alpha,\beta}^{A} C_{\alpha i}C_{\beta\alpha}(\alpha|\vec{r}_A|\beta) \quad (18)$$

Although the second term, the polarization term, may seem inconsistent with the CNDO scheme, it is theoretically justified if one assumes that the ZDO approximation is only appropriate over spherically symmetric operators and is empirically justified for the superior results obtained by its inclusion.

In general, the use of Eq. (18) to estimate oscillator strengths yields reasonable results for weak transitions and overestimates the strong transitions by a factor of two or three. Part of the reason for this overestimate is the limitation on most calculations to a singles-only CI: higher excitations usually reduce calculated oscillator strengths by reducing the weight of both ψ_o in the ground state description, and ψ_i^a, the principle component of the single excitation (see Eq. (18)). Part of the reason also resides in the evaluation of the polarization terms of Eq. (18) using Slater-type orbitals rather than the "better" orbitals parameterized on the energy.

For molecules composed of hydrogen and first row elements, CNDO/S and INDO/S execute in roughly the same time and the results obtained are roughly equivalent. For complexes containing transition metals the INDO/S method is somewhat slower, but this is a small price to pay for vastly improved results.[22]

There is no question that the CNDO/S and INDO/S theories, when applied to systems within their domain, yield useful results in interpreting spectroscopic and photochemical information. A survey of some results that are obtained are presented here as examples and indicate the wide range of information that can be obtained.

B. Some Results

The results obtained for the pi electron spectra of hydrocarbons is equally as impressive as those obtained from the PPP theory. An example is given for the triplet states of benzene in Table I, where INDO/S results are compared with experiment, and with those obtained from the *ab initio* calculations of Hay and Shavitt.[23] The most striking differences occur for the $^3B_{2u}(\pi-\pi^*)$ band, where the INDO/S results are 0.16 eV below the experimental value, and the *ab initio* results 1.40 eV above.

TABLE I. Benzene triplet states (eV).

State	Exp.	INDO/S[a]	*Ab initio*[b] Singles	Singles and Doubles	Singles and Doubles and Triples
$^3B_{1u}$	3.9	3.90	3.67	5.20	3.83
$^3E_{1u}$	4.7	4.8	5.15	5.78	4.98
$^3B_{2u}$	5.6	5.44	6.01	7.76	7.00
$^3E_{2g}$	6.6	7.06	7.86	8.59	7.28

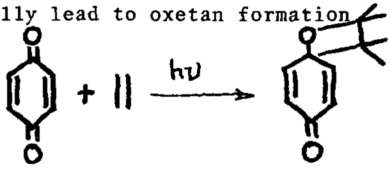
a) Reference 14

b) Reference 23

Table II presents results obtained for the singlet spectrum of pyridine.[13] The numerical agreement with experiment is striking. These results are particularly interesting in their prediction of a second (n,π^*) transition at 44,000 cm^{-1}, midway between two weakly allowed transitions. This prediction of something new in an "old" spectrum has been confirmed.

Figure 2 shows the calculated dependence of the spectrum of pyridazine (1,2-diazine) on the N-N bond length. The geometry of pyridazine is somewhat uncertain. The best agreement with the experimental spectrum comes from a choice of N-N bond length of 1.32 Å, which is in good accord with the N-N length of 1.321 Å found in s-tetrazine.[14]

Figure 3 shows the dependence of the triplet states of benzoquinones on methyl ring substitution. The assignment of the lowest triplet in parabenzoquinones is $^3B_{1g}(n,\pi^*)$ while that for duroquinone (tetramethyl parabenzoquinone) is $^3B_{1g}(\pi,\pi^*)$. The electron densities of these states are consistent with the hypothesis that $^3(n,\pi^*)$ states photochemically lead to oxetan formation



while $^3(\pi,\pi^*)$ states lead to cyclobutanes

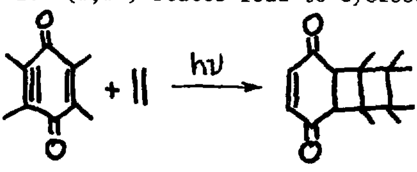


TABLE II. Singlet states of pyridine (1000 cm^{-1}).

Calculated Type	Calculated Energy (Osc.)	Observed
$^1B_1(n, \pi^*)$	34.7 (0.01)	34.8 (0.003)
$^1B_2(\pi, \pi^*)$	38.6 (0.07)	38.4 (0.04)
$^1A_2(n, \pi^*)$	44.0 (forb.)	
$^1A_1(\pi, \pi^*)$	49.7 (0.06)	49.8 (0.10)
$^1A_1(\pi, \pi^*)$	56.9 (0.91)	55.0 (1.30)
$^1B_2(\pi, \pi^*)$	56.9 (0.88)	
$^1A_2(\pi, \sigma^*)$	59.4 (forb.)	56.4 (diffuse)?
$^1B_1(n, \pi^*)$	61.8 (0.01)	
$^1B_2(\pi, \pi^*)$	62.7 (0.01)	

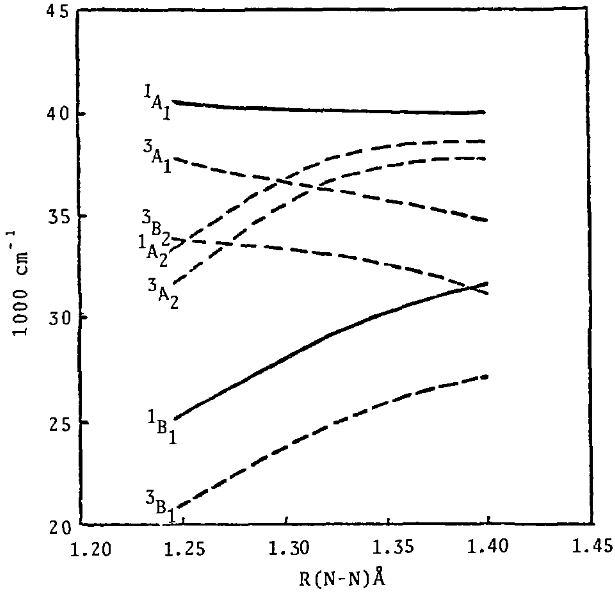


Fig. 2. The spectrum of pyridazine (1-2 diazene) as a function of the N-N bond length, from Ref. 14. Heavier lines are used to designate greater intensity, dashed lines represent forbidden transitions.

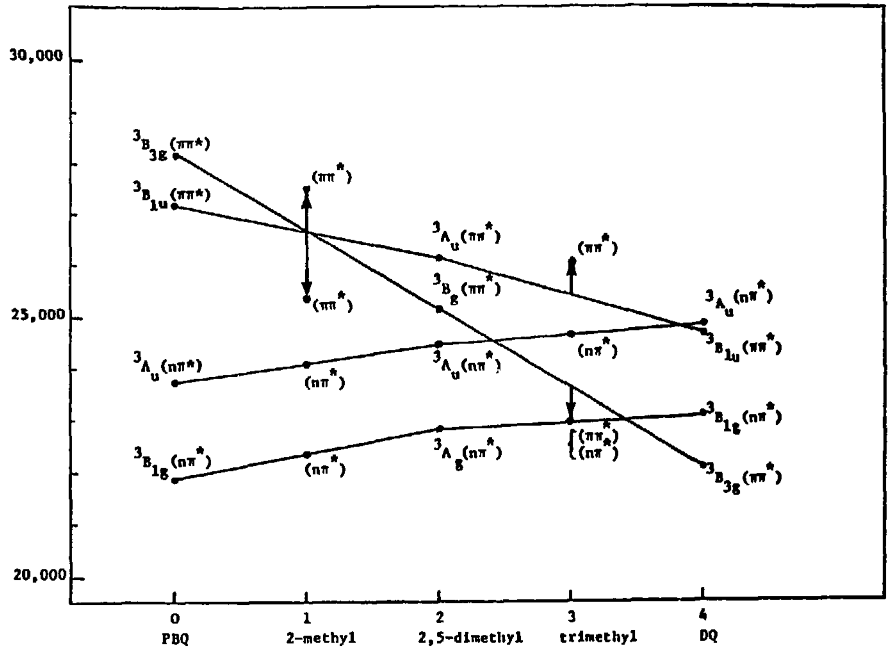


Fig. 3. The triplet states of parabenzoquinones (PBQ) vs. methyl substitution. DQ = duroquinone = tetramethyl PBQ. Taken from Ref. 24.

These results contrast with conclusions drawn previously about the nature of the excited states of these compounds through emission studies that suggest the same photochemistry regardless of methyl substitution.[24]

Table III presents the results of calculations on the excited states of ferrocene.[20] The results are striking. The detailed assignment obtained via INDO/S for band II differs from both *ab initio* CI and ΔE(SCF) results[25] (although the latter are nearly degenerate). Examination of the vibrational structure (similar to that observed for band III), however, suggest the INDO/S order of these states is correct. Bands I-III are (d,d^*); at higher energies the excitations are of charge-transfer tyhe. The INDO/S results are the only ones that show any *a priori* predictive strength for these charge transfer excitations.[20]

Table IV shows the calculated spectrum of $CuCl_2$ obtained from many different techniques. The conclusions of this work are that two (d,d^*) bands split by ~2000 cm^{-1} exist within the structure, with $\gamma_{max} \approx 9000$ cm^{-1}, and that two nearly degenerate charge-transfer excitations are reponsible for the maximum at 19,000 cm^{-1}. Although all the methods presented in this table could be used to reach this conclusion (except the SWX_α results), the INDO/S results, obtained in this case via UHF ΔE(SCF) calculations, are the only ones which yield such very good numerical agreement.[26]

The above are just a few examples of applications of

TABLE III. Assignment of the ferrocene spectrum (energy in units of 1000 cm^{-1}) (from Ref. 20).

	Observed [a]			INDO/S [b]			*Ab initio* [c]		MSX$_\alpha$ [d]
							CI	ΔE(SCF)	
I.	18.9		$^3E_{1g}(^3E_1'')$	20.5		$^3E_1''$	14.6		
				20.6		$^3E_1''$	to		
				20.6		$^3E_2''$	15.1		
II.[h] a)	21.8		$^1E_{1g}(^1E_1'')$	21.7		$^1E_1''$	26.7	14.2	20.5
b)	24.0		$^1E_{2g}(^1E_2'')$	23.9		$^1E_2''$	21.2	13.3	25.2
III.	30.8		$^1E_{1g}(^1E_1'')$	31.9		$^1E_1''$	46.3	21.8	25.2
IV.	37.7	(0.02)[e]		36.9		$^1E_2'$	All		36.5[f]
				39.7		$^1A_1'$	above		
				39.9		$^1A_2'$	73.0		
V.	41.7	(0.01)		41.2	(0.06)	$^1E_1'$			43.9[f]
	42.2	(0.01)		42.4		$^1A_1''$		60.7	36.5[f]
				42.9		$^1E_1''$		61.4	
				44.7		$^1E_2'$			39.2[f]
VI.	46.9	(0.10)[e]	$^1A_{2u}(^1A_2'')$	45.3	(0.29)	$^1A_2''$		62.0[g]	36.5[f]
				45.9		$^1E_2''$		46.3[g]	
				47.7		$^1A_1'$			
	49.7[e]			50.2	(0.00)	$^1E_1'$			47.1
VII.	50.9	(0.69)[e]	$^1A_{2u}(^1A_2'')$	50.3	(0.02)	$^1A_2''$			43.9[f]
	53.1			52.3	(0.03)	$^1E_1'$			

Footnotes to Table III

a. Y.S.Sohn, D.C.Hendrickson and H.B.Gray, J. Am. Chem. Soc. 93, 3603 (1971).

b. On eclipsed ferrocene. The group theoretical assignments have been made by correspondence between D_{5n} and D_{5d}, from Ref. 20.

c. On staggered ferrocene, Ref. 25.

d. Average singlet-triplet values from N. Rosch and K.H.Johnson, Chem. Phys. Letters 24, 179 (1974).

e. A.T.Armstrong, F.Smith, E.Elder, and S.P.McGlynn, J. Chem. Phys. 46, 4321 (1967).

f. The irreducible representations that result from orbital excitations have not been separated.

g. See text for discussion of these states. They are arranged in this table only according to symmetry type, not orbital character.

h. II is a transition centered at 22,000 cm^{-1} but analyzed in two transitions, IIa and IIb.

TABLE IV. Excitation energies (cm^{-1}) from $^2\Sigma_u^+$ ground state of $CuCl_2$.[a]

METHOD	$^2\Pi_g$	$^2\Delta_g$	$^2\Pi_u$	$^2\Sigma_u^+$
UHF[a)]	1,575	3 818	39 191	39 264
RHF[a)]	1,616	3 600	-	-
RHF[a),b)]	1,400	3 600	-	-
RHF[c)]	2,310	4 753	-	-
RHF-CI[c)]	8,230	10 482	15 230	-
MSX_α-TS[a)]	6,198	23 961	16 673	-
INDO[a),d)]	6,500	8 738	26 070	29 000
EH [e)]	3,550	4 275	22 750	22 500
EH [f)]	5,485	7 100	-	-
Exp. (j)	9,000[g)]	18 000[g)]	-	-
Exp. (k)	4,000[h)]	9 000[g)]	19 000[g)]	50 000[g)]
Exp. (l)	9,000[g)]	9 000 - 19 000	19 000[g)]	44 800[g)]
Exp. (m)	4,200[i)]	10 800[g)]		

a) Reference 26

b) Broken Symmetry

c) C.D. Garnier, I.H. Hillier and C. Wood, Inorg.Chem. 17, 168 (1978).

d) R(Cu-Cℓ) = 2.17A (geometry optimization).

e) R(Cu-Cℓ) = 2.37A

f) R(Cu-Cℓ) = 2.20A

g) observed peaks

h) estimated from crystal field theory

i) estimated from angular overlap model

j) G.E. Leroi, T.C. James, J.T. Hougen and W. Klemperer, J.Chem.Phys. 36, 1879 (1962).

k) C.W. DeKock, and D.M. Gruen, J.Chem.Phys. 44, 4387 (1966).

l) D.W. Smith, Chem.Phys. Letters 6, 83 (1970).

m) A.B.P. Lever and B.R. Hollebone, Inorg.Chem. 11, 2183 (1972).

the INDO/S technique, and are far from exhaustive or even representative of the great number of applications that have been made.

C. Discussion

For molecules examined within the "domain of applicability", the CNDO/S and INDO/S methods are remarkably accurate. In a survey of over 1000 experimental bands, the standard deviation between calculated and experimental results is $\pm 1000\ cm^{-1}$ for the singlets, and just slightly greater for the triplets (where less experimental information is available). Although this comparison is not quite balanced, as it weights allowed bands where greater accuracy is expected more heavily than forbidden bands, it is indicative of a reliable theory. Occasionally calculations of transitions have been off by $4000\ cm^{-1}$ or even more, and reversals between calculated and experimental transitions have occurred. But even these modest deviations cause few difficulties in interpreting the spectra if the results are taken in conjunction with experiment. I, personally, do not trust the absolute results of any calculations, either *ab initio* or semi-empirical, without carefully looking over my shoulder at the experiment. This is, I think, all one can expect from any technique that is not an exact application of an exact theory.

The "domain of applicability" of the methods described in these notes set up, in a certain sense, the ground rules for the systems that can be examined. The restrictions are generally of two types. The first is that the systems should not be too small, where the Rydberg states are important in describing the low energy states of interest; the latter, that the system should not be so large that as a consequence double or higher excitations are major contributors to the low energy states.

The incorporation of Rydberg orbitals within ZDO models is an important extension of the method and has been examined by several investigators.[27,28] These methods generally increase the basis set to include 2s,2p orbitals on hydrogen and 3s,3p orbitals on the first row elements. This modification has not been as successful as one might first suspect. Part of the explanation for this appears to be in the fact that the parameterization allows too much Rydberg mixing into the ground state description, lowering ground state energy and damaging the valence spectra agreement. Another reason may be the failure to include 3d polarization functions in some systems. These shortcomings, and the reported increased computing times, all suggest the use of a pseudo-potential to create the Rydberg orbitals after the SCF step, but before the CI.[29] In addition, it is probably unnecessary to include 3s orbitals, for example, on all first row elements of the system, providing all necessary symmetry representations can be created by those that are included.

For large systems, or for systems with two or more transition elements, it may be necessary to include higher excitations than the singles that were used in the parameterization of the method. This is a natural consequence of the fact that no amount of parameterization can account for the effect of a higher excitation if an actual higher excited configuration is in, or near, the region of spectroscopic interest. Although investigations with double or higher excitations often do lead to improved results, caution must be displayed in dealing with what correlation is included implicitly and what is included explicitly. Consistent would be a re-parameterization on a theory of all singles, doubles and triples (triples = doubly excited with respect to all singles), but such a large CI would be impossible for the systems of interest. More feasible would be a re-examination of the theoretical justification for such model Hamiltonians[30,31] and, perhaps, attempts to fold back higher excitations via perturbation theory.[32]

IV. THE LOCALIZED BOND MODEL

A. Theory

For many problems in quantum chemistry the only quantity of interest is the system energy. When this is the case other methods certainly become competitive with the Hartree-Fock SCF-MO theory. For example, one can start with any set of orbitals, including Schmidt orthogonalized atomic orbitals, form expansion determinants over these orbitals, Eq. (9), and minimize the molecular energy with respect to these linear expansion coefficients, Eq. (10), never visiting the repeated SCF equations, and very quickly not only obtain a better (lower) energy than the SCF-MO energy for that basis, but also the Hartree-Fock limit (best single determinant). The reason for this is, of course, well known, as one explicitly includes electron correlation in such a determinant expansion. What, perhaps, is less appreciated is that this procedure often proceeds more rapidly than the SCF-MO procedure. Although the energy of this direct configuration interaction (DCI) process is often good over large ranges of the potential energy surface (with sufficient care), the wave function obtained may be "unbalanced" and not very useful in itself.

Recently Diner, Malrieu and Claverie introduced a series of approximations within the framework of DCI and created a consistent scheme for calculating molecular energy.[33] Their model rests upon treating the CI matrix by perturbation theory. The technique is considerably simplified by adopting the CNDO approximation of Pople and co-workers.[6] The result is a scheme which executes five to ten times more rapidly than its SCF counterpart and produces an energy that is reliable (at least conceptually) over large regions of the potential energy surface. I quickly outline their model, perturbation configuration interaction localized orbitals (PCILO) here, and then introduce onto this scheme a variational procedure (called

PVCILO), and conclude with a comparison of results and a discussion of general utility. Since the CNDO parameters are derived to match minimum basis set *ab initio* SCF results, the PCILO model is one in which all the correlation energy is to be included explicitly. This of course does not preclude a re-parameterization directly on experiment.

(1) One begins the PCILO procedure by assuming a minimum basis set of valence-type orbitals. The neglect of the inner shell is accommodated by parameterizing one-center terms from experimental atomic information, Eq. (11*B*), and utilizing a scaled nuclear-electronic attraction term, Eq. (11*D*).

(2) One hybridizes the basis. The exact form of the hybridization does not appear to greatly alter the results. The method suggested is one that maximizes the overlap according to Del Re's procedure.[34] Hybridization has many advantages. Primary among these is the fact that although overlap between "bonding" hybrids is large, that between nonbonding hybrids is small, and generally less than 0.25. Another advantage is that atomic parameters chosen for molecules seem more transferable from molecule to molecule than the corresponding parameters for the primitive S and P functions.[35,36]

(3) From the hybridized set $\{X_i\}$ are formed localized bonds and antibonds $\{\phi_i\}$ between adjacent atoms

$$\phi_i = \cos\alpha\, X_i + \sin\alpha\, X_{i+1}$$

$$\phi_i = -\sin\alpha\, X_i + \cos\alpha\, X_{i+1}$$

For simplicity α is set to 45°, although PCILO computer programs allow for α to be chosen variationally. Choosing such a set of bonding orbitals then completes the rationale for invoking the CNDO approximation. The local nature of the anti-bonds, confined more tightly to the molecule than, say, virtual molecular orbitals, provides a particularly convenient set for determinantly expansions of the wavefunction.[37,38]

(4) The CNDO approximations are now invoked to simplify the calculation. This is not an essential part of the theory, but is of course, desirable for investigating large systems. The choice of a basis of bonds and antibonds between hybrids, however, makes the zero differential overlap idea (or at least the neglect of differential diatomic overlap idea) much sounder than in the SCF-MO theory, where overlap between atomic orbitals often exceeds 0.50. The CNDO approximation reduces the number of integrals from N^4 to N^2.

(5) The zero-order wavefunction ψ_0 is formed from an antisymmetrized product of doubly occupied bonds. Upon the fact that this is most often a good starting point rests most of the current rationale of organic chemistry. It is here that the philosophy of the method enters. One admits from the start that the localized bond description will not provide the best single determinant of the system, which by definition is the Hartree-Fock description. But we fully intend, from

the start, to be beyond the single determinant and include correlation. In return one avoids the repeated N^3 matrix diagonalizations needed to solve Eq. (5). The zero-order energy is then given by

$$\varepsilon_0 + \varepsilon_1 = \langle\psi_0|H|\psi_0\rangle \equiv \langle 0|H|0\rangle$$

and its evaluation is proportional to the number of integrals, in this case N^2, rather than the complex N^3 step suggested in Eq. (5). $\langle 0|H|0\rangle$ is generally within 1-2% of E_{HF}.

(6) The wavefunction Ψ is approximated as ψ_0 plus contributions from determinants created by exciting one or more electrons from bonds to antibonds. The energy is then evaluated by third-order Rayleigh-Schrödinger perturbation theory.

$$E \approx \varepsilon_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \tag{19}$$

$$\varepsilon_2 = \sum_I \frac{\langle 0|V|I\rangle\ \langle I|V|0\rangle}{E_i - E_0} \tag{20}$$

$$\varepsilon_3 = \sum_{I,J} \frac{\langle 0|V|I\rangle\ \langle I|V|J\rangle\ \langle J|V|0\rangle}{(E_I - E_0)(E_J - E_0)} - \varepsilon_1 \sum_I \frac{\langle 0|V|I\rangle^2}{(E_I - E_0)^2} \tag{21}$$

The Epstein[39]-Nesbet[40] partitioning of the Hamiltonian has been utilized by Malrieu, Claverie and Diner[33] where, in this case

$$H = H_o + V \tag{22}$$

$$E_I = \langle I|H|I\rangle \tag{23}$$

$$\langle I|V|I\rangle = 0 \tag{24}$$

$$\langle I|H_o|J\rangle = 0 \ , \qquad I \neq J \tag{25}$$

A particular consequence of Eq. (24) is that $\varepsilon_1 = \langle 0|V|0\rangle = 0$.

Our version of PCILO follows that outlined above, except that the expansion is over proper spin states, whereas the originally disclosed version is over determinants. The principle consequence of this is that we include in second order, ε_2, large terms not included until third order, ε_3, when using a basis of determinants. Similarly, we include at third order terms which otherwise are neglected. This results in an expansion that seems more convergent. Numerical evidence indicates that $\gamma_2 \equiv \varepsilon_3/\varepsilon_2$ for proper spin states is approximately one-half that for determinants.[41]

Equation (14) is of interest for several reasons. The only N^3 step, where N is the number of basis functions (bonds), is the evaluation of the third-order energy, and this is a considerably easier N^3 step than the repeated matrix diagonalization of SCF-MO theory. Second, experience with the method has shown that the energy expression to third order is a reasonable representation of the system energy. These advantages are also chief disadvantages. There is still an N^3 step, making the method slower than some reliable semi-classical methods, although the latter suffer from "personal interpretations" of bonding. Second, the third-order energy expression yields results for some molecules below the exact energy, for others above, making some comparisons unreliable. Indeed, we have found that Eq. (19) can give energies that start above the exact energy, and as a bond stretches (or shrinks) slip below. Such a result shifts calculated minima of the potential energy surface.

A variational procedure can be obtained from the third order energy via the procedure of Goldhammer and Feenberg.[43,44] Assuming

$$\phi = \psi_0 + \lambda\psi'$$

where ψ' is the first order correction

$$\psi' = \sum_I \frac{\langle 0|V|I\rangle}{E_0 - E_I} \psi_I$$

and utilizing the variational principle we have

$$E < W \equiv \frac{\langle\phi|H|\phi\rangle}{\langle\phi|\phi\rangle} = \varepsilon_0 + \frac{\varepsilon_1 - (\lambda-1)^2\varepsilon_2 + \varepsilon_2 + \lambda^2\varepsilon_3}{1 - \lambda^2 S} \qquad (26)$$

where

$$S = \langle\psi'|\psi'\rangle = \sum_{I\neq 0} \frac{\langle 0|V|I\rangle\ \langle I|V|0\rangle}{(E_0 - E_I)} \qquad (27)$$

Minimizing Eq. (26) with respect to λ yields

$$E \leq W = \varepsilon_0 + \varepsilon_1 + \lambda\varepsilon_2$$

$$\lambda = \frac{1}{2S}\{(\gamma_1 - 1) + [(\gamma_1 - 1)^2 + 4S]^{\frac{1}{2}}\} \qquad (28)$$

$$\gamma_1 = \frac{(\varepsilon_3 - S\varepsilon_1)}{\varepsilon_2}$$

For PCILO with the Epstein-Nesbet partitioning,

$$\varepsilon_1 = 0 \quad , \qquad E \leq W = \varepsilon_0 + \lambda\varepsilon_2$$

$$\lambda = \frac{1}{2S}\{(\gamma_2 - 1) + [(\gamma_2 - 1)^2 + 4S]^{\frac{1}{2}}\} \quad , \qquad (28')$$

$$\gamma_2 = \frac{\varepsilon_3}{\varepsilon_2} \quad ,$$

The quantities that occur in Eq. (28') are calculated at the same time as in the third-order energy. It is no more difficult (or easy) to evaluate Eq. (28') than to evaluate Eq. (19). The N^3 step is still present in the evaluation of λ. Bartlett and Brandas have examined in some detail the utilization of such a variational perturbation approach and have related this scheme with others.[45]

The use of the perturbation-variation technique of Eq. (28') has several advantages and one important disadvantage. Among the advantages is the recognition that changes in λ, the variational parameter, with geometry change q, do not effect the energy to first order

$$\frac{dE}{dq} = \frac{\partial E}{\partial q} + \frac{\partial E}{\partial \lambda} \frac{\partial \lambda}{\partial q}$$

since

$$\frac{\partial E}{\partial \lambda} = 0$$

The suggestion then is to evaluate λ at one point of interest on the potential energy surface with an N^3 step, then probe the surface with fixed λ, λ^f, an N^2 step. Other regions of interest may come into focus in this cursory examination. λ is then evaluated at one point in this new region and the neighboring points corrected with a new λ^f. A third-order method then becomes proportional, chiefly to N^2, and as such competitive to N^2 semi-classical methods used to evaluate molecular conformation. A second advantage to the variation procedure is that even if the perturbation theory is not well defined (slowly convergent or even divergent), Eq. (28') still yields a useful bound. Table V shows an example of this for CH_3CN and CH_3NC, where the perturbation expansion for CH_3CN is creeping ($\varepsilon_2 < 0$, $\varepsilon_3 < 0$) and that for CH_3NC oscillating ($\varepsilon_2 < 0$, $\varepsilon_3 > 0$).

A major disadvantage of the variational procedure is that it tends to smooth out features of the potential energy surface, and the theory is not "size consistent"; that is, by insisting on the comforts of a bound, terms in the energy are introduced that do not grow properly with the size of the system. An example of this is given for two benzene molecules in Table VI. How important size inconsistency is depends on

TABLE V. Comparison of acetonitrile (CH_3CN) and methylisocyanide (CH_3NC).

CH_3CN	E(kcal/mole) R_{CN} = 1.46	R_{CN}(Å)	k_{CN}(mdynes/Å)
$E_0 + E_2$	-17063.9	1.439	15.3
$E_0 + E_2 + E_3$	-17557.6	1.431	16.8
$E_0 + \lambda^f E_2$	-17548.3	1.459	16.9
$E_0 + \lambda E_2$	-17548.3	1.456	17.0
E(SD)[a]	-17553.5	1.455	17.0

CH_3NC	E(kcal/mole) R_{CN} = 1.41	R_{CN}(Å)	k_{CN}(mdynes/Å)	ΔE[b](kcal/mole)
$E_0 + E_2$	-17777.5	1.413	15.3	-173.6
$E_0 + E_2 + E_3$	-17164.8	1.397	17.5	392.8
$E_0 + \lambda^f E_3$	-17434.8	1.423	18.3	113.5
$E_0 + \lambda E_2$	-17433.9	1.420	18.3	114.4
E(SD)[a]	-17455.1	1.420	18.0	98.4

a) From a diagonalization of the CI matrix of all single and double excitations.

b) $\Delta E = E(H_3CNC) - E(H_3CCN)$.

TABLE VI. PCILO and PVCILO calculations on the parallel plate dimer of benzene (kcal/mole).*

	$E_o+E_2+E_3$	$E_o+\lambda E_2$
2 × monomer energy	-59308.3	-59008.8
Dimer 10Å separated	-59308.3	-58905.6
Difference	0.0	-103.2 (0.16%)

*Nesbet-Epstein partitioning.

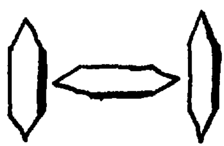
the nature of the problem. The perturbation theory results might be preferred whenever they do not yield results greatly different from those of the perturbation-variation procedure. When they do differ, this is generally an indication of a badly behaved perturbation sequence, and the variational results might be preferred as still yielding a useful bound.

B. Some Results

Although the PCILO technique has been put to many uses,[46] by far the most common and most successful are those that deal with the molecular conformation of very large systems, where only semi-classical methods can compete. Notable applications include studies on the conformation of peptides,[47] nucleic acids,[48] and phospholipids.[49] The method has also been applied to drug design, where structure-activity relations are crucial.[50] In addition PCILO has been used as a preliminary method to uncover minima in a potential energy surface in conjunction with subsequent *ab initio* calculations.[51]

In general, the results from the PCILO and PVCILO methods described mimic the results of the CNDO-SCF theory, but execute some ten to twenty times faster on the computer. As such, its failures might be expected to be those of the CNDO-SCF method. Most notable among those failures is the systematic stability given strained cycles with respect to the linear isomers. The PCILO model will contain some of the correlation energy, but for most of the applications reported the correlation correction does not seem to greatly alter results from what one would expect from the SCF counterpart. Notable exceptions deal with weakly bonded dimers where the localized zero-order description might be expected to be superior to the canonical "super" molecular orbital description, and more importantly downgrades, to some extent, the interest the CNDO approximations have for overbinding such systems. A comparison of PCILO results with those of CNDO and extended Hückel theory show the former far superior in describing the interaction between tetracyanoethylene and benzene.[52] In addition, John Cullen and I have been able to explain the herringbone structure of liquid benzene from the relative stabilities of

and



triplets over all others.[53] In this case the CNDO-SCF method overbinds by an impressive amount.[54] Care must be exercised in the interpretation of such results, however, as higher order perturbation theories must converge to the same result as SCF plus full CI, and the latter is likely to be similar to the original SCF results because of inherent errors in the CNDO approximation itself.[46,55] Very recently Lochmann and Holza[56] have examined a great many van der Waals complexes using the PCILO model with good success for those systems not possessing lone pairs (see discussion below).

C. Discussion

The PCILO method is interesting both from the applied point of view, where it provides an easy way to evaluate molecular energy and thus molecular conformation, and from the theoretical point of view, where it provides an easy model Hamiltonian and an interesting reference state of doubly occupied bonding functions that relate well with "classical" chemical concepts.

From the applied point of view the PCILO model is bound to fail whenever the corresponding CNDO-SCF method fails, for it has been parameterized on the latter, and most features present for strong bonding are present in the SCF theory. Notable exceptions, at least at third-order in perturbation theory, are found in examining weakly bound van der Waal or charge transfer complexes in which PCILO yields more reasonable results.

An obvious improvement of the model would be its extension to include INDO or NDDO integrals. The former has been recently accomplished by Douady, Barone, Ellinger and Subin, with improved results, especially for calculated angles and rotational barriers, as expected.[57,58] Part of this improvement, though, may be due to an improved description of the lone pairs as well as the inclusion of the additional one-center integrals of Eq. (17).

It is indeed tempting to apply the PCILO idea in an *ab initio* fashion.[59] This unfortunately is difficult, for the localized bonds are not orthogonal, thus making the integrals of Eqs. (19)-(28) difficult to evaluate. Orthogonalization of these bonds to one another creates delocalization of these bonds and re-introduces the integral transformation problem. The loss of Brillouin's Theorem (that single excitations interact with the reference configuration) that accompanies the localized bond description may be too high a price to pay if one must also transform integrals over bonds to integrals over the entire system. However, theories that treat the non-orthogonality problem – or corrections to the NDDO approximation – as a second perturbation, may show promise.

From the theoretical point of view the PCILO model has been extended to infinite order in single and double excitations, and to fourth order in singles, doubles, triples, and quadruples.[60] Although the former beyond fourth-order proceeds as N^5, and is thus slow in application, the fourth-order fully linked

TABLE VII. Localized bond study on ethylene (kcal/mole).

HF = Hartree Fock	-10708.2		
HF + SDCI[a]	-10750.7		
HF + SDCI + $\Delta\varepsilon_2$	-10752.5		
Localized Bond	Mollet-Plesset Determinant	Nesbet-Epstein Determinant	Nesbet-Epstein Spin States
LB = $\varepsilon_0 + \varepsilon_1$	-10672.7	-10672.7	-10672.7
LB + ε_2(singles)	-10697.9	-10706.2	-10705.9
LB + $\varepsilon_2 + \varepsilon_3$	-10749.9	-10753.0	-10754.9
LB + $\varepsilon_2 + \varepsilon_3 + \varepsilon_4$ (SD)	-10756.7	-10765.6	-10765.3
LB + $\varepsilon_2 + \varepsilon_3 + \varepsilon_4$ (SDTQ)[b]	-10757.0	-10765.6	-10763.9
Padé on 4th order	-10763.1		
LB + SD-MBPT[c]	-10764.2		

Time (sec.) Amdahl 450 V5	
HF	3.2
CI	32.6
HF + CI	35.8
LB + SDTQ - 4th Order	0.3

a) SDCI = all singles and doubles CI

b) SDTQ = singles, doubles, triples and quads at 4th order

c) The Padé Approximate at 6th order and the direct summation at 8th order are generally within ±0.1 kcal/mole of converged result.

correct order of states from our interpretation of the experimental data (in contrast to the data on ferrocene using a smaller basis set, Table III), the transition energies are an average factor of 1.7 too small. The INDO/S results for these copper complexes were usually quite close to the experimental values.

Turning now to the PCILO method – quite simply few laboratories would be able to examine the geometric conformations of, say, valinomycine,[61] and those that could should certainly have made a preliminary investigation using PCILO and PVCILO.

REFERENCES

1. C.C.J.Roothaan, Rev. Mod. Phys. 23, 69 (1951); ibid. 32, 179 (1960).

2. See for example, H.E.Popkie and J.J.Kaufman, Int. J. Quantum Chem., Quant. Biol. Symp. 2, 279 (1975); Int. J. Quantum Chem. S10, 47 (1976).

3. T.A.Halgren, D.A.Kleier, J.H.Hall,Jr., L.D.Brown, and W.N.Lipscomb, J. Am. Chem. Soc. 100, 6595 (1978).

4. R.G.Parr, The Quantum Theory of Molecular Electronic Structure (W. A. Benjamin, Inc., 1964).

5. I.Fischer-Hjalmers and M.Sundbom, Acta. Chem. Scand. 22, 607 (1968); B.Roos and M.Sundbom, J. Molec. Spectros. 36, 8 (1970).

6. J.A.Pople, D.P.Santry and G.A.Segal, J. Chem. Phys. 43, S129 (1965); J.A.Pople and G.A.Segal, J. Chem. Phys. 43,

S136 (1965); ibid. 44, 2389 (1966).

7. See for example, G.Karlsson and M.C.Zerner, Int. J. Quantum Chem. VII, 35 (1973).

8. J.Del Bene and H.H.Jaffé, J. Chem. Phys. 48, 1807 (1968); ibid. 4050 (1968).

9. R.Pariser, J. Chem. Phys. 21, 568 (1953).

10. N.Mataga and K.Nishimoto, Z. Physik Chem.(Frankfurt) 13, 140 (1957).

11. K.Ohno, Theoret. Chim. Acta 3, 219 (1964).

12. G.Klopman, J. Am. Chem. Soc. 86, 4550 (1964).

13. J.E.Ridley and M.C.Zerner, Theoret. Chim. Acta 32, 111 (1973).

14. J.E.Ridley and M.C.Zerner, Theoret. Chim. Acta 42, 223 (1976).

15. J.A.Pople, D.L.Beveridge and P.A.Dobosh, J. Chem. Phys. 47, 2026 (1967).

16. J.C.Slater, Quantum Theory of Atomic Structure, Vol. I (McGraw-Hill, New York, 1960).

17. D.W.Clack, Mol. Phys. 27, 1513 (1974);
D.W.Clack, N.S.Hush and J.R.Yandle, J. Chem. Phys. 57, 3503 (1972).

18. W.T.A.M.van der Lugt, Int. J. Quantum Chem. 6, 859 (1972).

19. H.P.Figeys, P.Geerling, and C.van Alsevoy, Int. J. Quantum Chem. 11, 705 (1977).

20. M.C.Zerner, R.Kirchner, G.Lowe and U.T.Mueller-Westerhof, submitted for publication.

21. H.M.Chong, H.H.Jaffé, C.A.Masmanidis, J. Phys. Chem. 79, 1118 (1975).

22. A.Bacon and M.C.Zerner, Theoret. Chim. Acta __, (1979).

23. P.J.Hay and I.Shavitt, J. Chem. Phys. 60, 2865 (1974).

24. N.J.Bunce, J.E.Ridley and M.C.Zerner, Theoret. Chim. Acta 45, 283 (1977).

25. M.M.Rohmer, A.Veillard and M.H.Wood, Chem. Phys. Lett. 29, 466 (1976).

26. P.Corrêa de Mello, M.Hehenberger, S.Larsson and M.C. Zerner, submitted for publication.

27. D.R.Salahub and C.Sandorfy, Theoret. Chim. Acta 20, 227 (1971).

28. W.Hague, J. Chem. Phys. 67, 3629 (1977).

29. J.Zinck and M.C.Zerner, unpublished work.

30. S.Iwata and K.Freed, J. Chem. Phys. 61, 1500 (1974); Chem. Phys. Lett. 28, 176 (1974).

31. B.H.Brandow, Int. J. Quantum Chem. XV, 207 (1979).

32. D.Edwards and M.C.Zerner, in progress.

33. S.J.Diner, J.P.Malrieu and P.Claverie, Theoret. Chim. Acta 13, 1 (1969); ibid. 13, 18 (1969);
S.Diner, J.P.Malrieu, F.Jordan and M.Gilbert, Theoret. Chim. Acta 15, 100 (1969).

34. G.Del Re, Theoret. Chim. Acta 1, 1 (1963).

35. D.B.Cook, P.C.Hollis, and R.McWeeny, Mol. Phys. 13, 553 (1967).

36. J.E.Eilers and D.R.Whitman, J. Am. Chem. Soc. 95, 2067 (1973).

37. O.Sinanoglu, Adv. Chem. Phys. 6, 315 (1964).

38. R.K.Nesbet, Adv. Chem. Phys. 9, 321 (1965).

39. P.S.Epstein, Phys. Rev. 28, 695 (1962).

40. R.K.Nesbet, Proc. Roy. Soc. A230, 312 (1955).

41. J.Cullen and M.C.Zerner, Int. J. Quantum Chem. 9, 343 (1975).

42. For example, H.A.Scheraga, in Advances in Physical Organic Chemistry, Vol. 6, V.Gold, editor (Academic Press, 1968), pp.103-184.

43. P.Goldhammer and F.Feenberg, Phys. Rev. 101, 1233 (1956).

44. P.Goscinski and E.Brändas, Chem. Phys. Lett. 2, 299 (1968); ibid., Phys. Rev. A1, 552 (1970); ibid., Int. J. Quantum Chem. 3S, 383 (1970).

45. R.J.Bartlett and E.J.Brändas, J. Chem. Phys. 56, 5467 (1972).

46. J.P.Malrieu, in Electronic Structure Calculations, Part A, G.A.Segal, editor (Plenum Press, 1977), pp.69-104.

47. B.Pullman and A.Pullman, Adv. Prot. Chem. 28, 347 (1974).

48. B.Pullman and A.Saron, Prog. Nucl. Acid Res. Mol. Biol.

49. B.Pullman and A.Saron, Int. J. Quantum Chem., Quant. Biol. Symp., in press.

50. B.Pullman, in Molecular and Quantum Pharmacology, E.Bergman and B.Pullman, editors (Reidel, 1974).

51. H.Berthod and N.Gresch, FEBS Lett. 53, 199 (1975).

52. R.Arnaud, D.Faramond-Baud, and M.Gelus, Theoret. Chim. Acta 31, 335 (1973).

53. J.Cullen, M.Sc. Thesis, Dept. of Chemistry, University of Guelph, Ontario, Canada (1976).

54. F.Grein and K.Weiss, Theoret. Chim. Acta 34, 315 (1974).

55. D.B.Chestnut and P.E.Wormer, Theoret. Chim. Acta 20, 250 (1971).

56. R.Lockman and P.Hobza, Int. J. Quantum Chem. XV, 73 (1979).

57. J.Douady, Y.Ellinger and R.Subra, Chem. Phys. Lett. 56, 38 (1978).

58. J.Douady, Y.Ellinger and R.Subra, Int. J. Quantum Chem., in press.

59. J.Cullen and M.C.Zerner, to be published.

60. B.Maigret and B.Pullman, Theoret. Chim. Acta 37, 17 (1975).

APPENDIX
PERTURBATION THEORY

Know

$$H_o\psi_\alpha^o = E_\alpha^o\ \psi_\alpha^o$$

want

$$H\psi = E\psi$$

$$H = H_o + V \quad , \qquad V < H$$

Consider

$$H_o|\psi_o^o> = E_o^o|\psi_o^o>$$

$$H|\psi> = E|\psi>$$

$$<\psi|H_o|\psi_o^o> = E_o^o <\psi|\psi_o^o> \qquad (a)$$

$$<\psi_o^o|H|\psi> = E<\psi_o^o|\psi> \qquad (b)$$

Assume "intermediate" normalization

$$<\psi_o^o|\psi> = 1$$

$$\Rightarrow \quad (b) - (a)^+ \Rightarrow (E-E_o^o) = \Delta E = <\psi_o|V|\psi>$$

$$\boxed{E = E_o + <\psi_o|V|\psi>}$$

$$(H_o + V)\psi = E\psi$$

$$(H_o+V+\varepsilon)\psi = (E+\varepsilon)\psi \qquad (\varepsilon \text{ is arbitrary shift})$$

$$(\varepsilon - H_o)\psi = (V - E + \varepsilon)\psi$$

$$\psi = \underbrace{(\varepsilon - H_o)^{-1}}_{\text{resolvent}} (V - E + \varepsilon)\psi$$

$$P\psi_o^o = \psi_o^o$$

$$P\tilde{\psi} = \psi_o^o \quad \text{or zero}$$

$$(P+Q)\tilde{\psi} = \tilde{\psi}$$

P and Q are projectors, i.e.,

$$P = |\psi_o^o\rangle \langle\psi_o^o|$$

$$P|\psi_o^o\rangle = |\psi_o^o\rangle \langle\psi_o^o|\psi_o^o\rangle = |\psi_o^o\rangle$$

$$Q = \sum_{\alpha \neq 0} |\psi_\alpha^o\rangle \langle\psi_\alpha^o|$$

1) $P^2 = P$, $Q^2 = Q$, $PQ = 0$

idempotent

mutually self-exclusive

2) $P+Q = 1$

$$\psi = (P+Q)\psi = \psi_o + Q\psi$$

$$Q\psi = Q(\varepsilon - H_o)^{-1} (V - E + \varepsilon)$$

$$\psi = \psi_o + Q(\varepsilon - H_o)^{-1} (V - E + \varepsilon)\psi$$

Iterate

$$\psi = \psi_o + Q(\varepsilon - H_o)^{-1} (V - E + \varepsilon)\psi_o + \ldots$$

$$= \sum_{n=0} [Q(\varepsilon - H_o)^{-1} (V - E + \varepsilon)]^n \psi_o$$

$$E = E_o + \langle\psi_o|V|\psi\rangle$$

$$= E_o + \sum_{n=0} \langle\psi_o|V[Q(\varepsilon - H_o)^{-1} (V - E + \varepsilon)]^n|\psi_o\rangle$$

$$= E_o + \sum_{i=0}^{\infty} \varepsilon_{i+1}$$

$$\varepsilon_{i+1} = \langle\psi_o|V[Q(\varepsilon - H_o)^{-1} (V - E + \varepsilon)]^i|\psi_o\rangle$$

$$\varepsilon_1 = \langle\psi_o|V|\psi_o\rangle$$

$$\varepsilon_2 = \langle\psi_o|VQ(\varepsilon - H_o)^{-1} (V - E + \varepsilon)|\psi_o\rangle$$

$$= \sum_{i} \langle\psi_\alpha|V|\psi_\alpha^o\rangle \langle\psi_\alpha^o|(\varepsilon - \varepsilon_\alpha^o)^{-1} V|\psi_o\rangle$$

BRILLOUIN-WIGNER PERTURBATION THEORY

$$\varepsilon = E$$

$$E = E_o + \sum_{n=0} \langle\psi_o|V[Q(E - H_o)^{-1} V]^n|\psi_o\rangle$$

RAYLEIGH SCHRÖDINGER PERTURBATION THEORY

$$\varepsilon = E_o^o$$

$$E = E_o + \sum_{n=0} \langle\psi_o|V[Q(E_o-H_o)^{-1}(V-E+E_o)]^n|\psi_o\rangle$$

$$-E+E_o = -\varepsilon_1 - \varepsilon_2 - \varepsilon_3 - \ldots$$

COMMON PARTITIONINGS:

$$\underset{\sim}{H}_{\underset{\sim}{C}} = \underset{\sim}{H}^{diag} + \underset{\sim}{V}$$

or

$$H = \sum_{\alpha} |\psi_\alpha^o\rangle \langle\psi_\alpha^o|H|\psi_\alpha^o\rangle \langle\psi_\alpha^o|$$

$$V = \sum_{\alpha\neq\beta} |\psi_\alpha^o\rangle \langle\psi_\alpha^o|H|\psi_\beta^o\rangle \langle\psi_\beta^o|$$

$$E_I = \langle\psi_I^o|H|\psi_I^o\rangle \equiv \langle I|H|I\rangle$$

$$\langle I|V|I\rangle = 0$$

$$\langle I|H_o|J\rangle = 0 \qquad (I \neq J)$$

$$\Rightarrow \varepsilon_1 = \langle 0|V|0\rangle = 0$$

EPSTEIN-NESBET

$$H_o = \sum_i f(i) \quad , \qquad f(i) = h(i) + V(i)$$

$$\varepsilon_o = \sum_\alpha \varepsilon_\alpha \quad , \qquad \varepsilon_\alpha = h_\alpha + \sum_\beta^{occ} \langle\alpha\beta\|\alpha\beta\rangle$$

if f(i) is a Fock type operator,

$$V = \sum_{i<j} \frac{1}{r_{ij}} - \sum_i V(i)$$

MØLLER-PLESSET PARTITIONING

$$<0|H_o+V|0> = \varepsilon_o^{MP} + \varepsilon_1^{MP}$$

$$= \sum_{\alpha}^{occ} \varepsilon_\alpha - \sum_{\alpha,\beta} (2J_{\alpha\beta} - K_{\alpha\beta}) = E_o^{NE}$$

Note that $\varepsilon_1 \neq 0$, but

$$E_I^A - E_o = \varepsilon_A - \varepsilon_I$$

$$E_{IJ}^{AB} - E_o = \varepsilon_A + \varepsilon_B - \varepsilon_I - \varepsilon_J$$

etc.

PREPARATION FOR CI

Form integrals over MO's

$$[ij\|k\ell] = \int \phi_i(1)\phi_j(1)\,\frac{1}{r_{12}}\,\phi_k(2)\phi_\ell(2)\;d\tau_1 d\tau_2$$

$$\phi_i = \sum_{p=1}^{N} c_{pi}\, f_p$$

$$\{pq\|rs\} = \int f_p(1)f_q(1)\,\frac{1}{r_{12}}\,f_r(2)f_s(2)\;d\tau_1 d\tau_2$$

$$[ij\|k\ell] = \sum_{pqrs} c_{pi}c_{qj}c_{rk}c_{s\ell}\,\{pq\|rs\}$$

N^q as written

Define

$$(\underset{\sim}{F}^{pq})_{rs} = \{pq\|rs\}$$

$$\underset{\sim}{G}^{pq} = \underbrace{\underset{\sim}{C}^T\,(\underbrace{\underset{\sim}{F}^{pq}\,\underset{\sim}{C}}_{N^3 \text{ for each } pq})}_{N^3 \text{ for each } pq} \qquad = 2N^5 \text{ multiples}$$

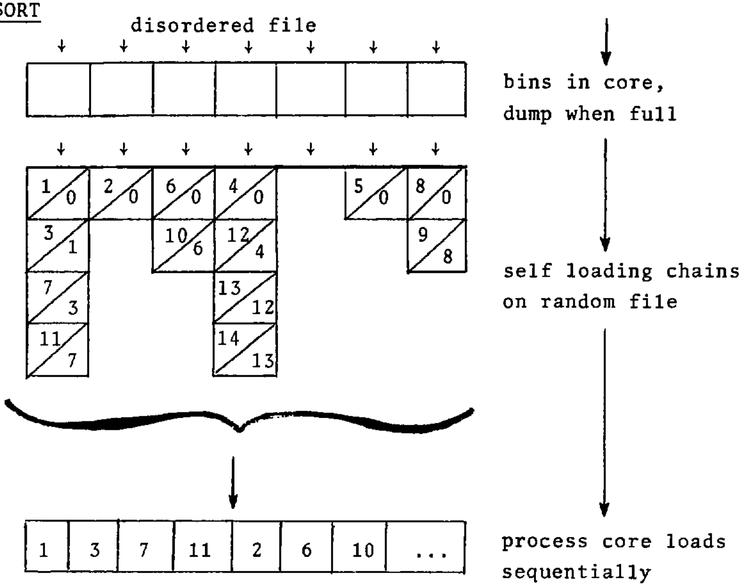
Define

$$(\underset{\sim}{M}^{k\ell})_{pq} = G^{pq}_{k\ell}$$

$$\{ij\|k\ell\} = \{\underset{\sim}{C}^T\,\underset{\sim}{M}^{k\ell}\,\underset{\sim}{C}\}_{ij} \qquad 2N^5 \text{ multiples}$$

so cost goes like N^5. Notice this process requires two sorting steps to arrange arrays in best order for next step.

SORT



COMPLICATIONS DUE TO SYMMETRY

$$[ij\|k\ell] = [ji\|k\ell] = [ij\|\ell k] = [ji\|\ell k]$$

$$[k\ell\|ij] = [k\ell\|ji] = [\ell k\|ij] = [\ell k\|ji]$$

Need only store $i \geq j \geq k \geq \ell$

$$\{[ij\|k\ell],\quad [ik\|j\ell],\quad [i\ell\|jk]\}\quad \text{"triplet"}$$

Point group symmetry: many integrals are zero, arrange in non-zero blocks (Γ_i = symmetry type)

$$[\Gamma_1\Gamma_1'; \Gamma_2\Gamma_2'; ij\|k\ell]$$

$$\Gamma_1 \otimes \Gamma_1' \otimes \Gamma_2 \otimes \Gamma_2' = \text{totally symmetric}$$

Consideration of symmetry reduces cost but complicates programs.

CONFIGURATION INTERACTION

Basic

Assume

$$\Psi \cong \sum C_I \Phi_I$$

Determine "best" C_I by minimizing $\langle\Psi|H|\Psi\rangle/\langle\Psi|\Psi\rangle$. There are several approaches:

H Matrix

$$H_{IJ} = \langle\Phi_I|H|\Phi_J\rangle$$

$$\sum_J H_{IJ}C_J = E\ C_I \qquad \text{(matrix eigenvalue)}$$

Small matrix methods: Jacobi, Givens, QR, LR, etc.

Large matrix methods: incomplete expansion

$$\underset{\sim}{C}^{(n)} = C^{(n-1)} + \sum_{i=1}^{k} \alpha_i\ \underset{\sim}{b}_i^{(n-1)}$$

Vary α_i to get best $\underset{\sim}{C}^{(n)}$. Select next $\underset{\sim}{b}_i$ set by some systematic scheme, so $\lim_{n\to\infty} \underset{\sim}{C}^{(n)} \longrightarrow \underset{\sim}{C}$. $\underset{\sim}{b}_i$ usually picked by

$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}$ ... or first order perturbation theory correction to $C^{(n-1)}$.

"Direct" CI

$$H_{IJ} = \sum_{\mathcal{J}} \mathcal{B}_{\mathcal{J}}^{IJ} \mathcal{J}$$

where $\mathcal{J}$ = integral and $\mathcal{B}$ = coefficients (mostly zero!)

$$\sum_{\mathcal{J}} \mathcal{J} \left(\sum_{J} \mathcal{B}_{\mathcal{J}}^{IJ} C_J \right) = E\ C_I$$

solved using large matrix and perturbation correction method. $\mathcal{B}$ *built into program logic or data statements.*

Semi-Direct CI

Many "H matrix" CI programs form $\mathcal{B}_{\mathcal{J}}^{IJ}$ as a "formula tape". Then one can solve

$$\sum_{\mathcal{J}} \mathcal{J} \left(\sum_{J} \mathcal{B}_{\mathcal{J}}^{IJ} C_J \right) = E\ C_I$$

bringing $\mathcal{J}$ and $\mathcal{B}$ from scratch files.

Specialized CI

Closed shell SCF + all single excitations.

Closed shell + all double excitations (self-consistent pairs).

CONSTRUCTION OF H

A. Integral driven

B. Formula driven

C. Partial formula tape

D. Complete formula tape

A. 1. Read in one $\mathcal{I}$

2. Determine all non-zero $\mathcal{G}_{\mathcal{I}}^{IJ}$ involving this $\mathcal{I}$

3. Store $(\mathcal{I}\mathcal{G}_{\mathcal{I}}^{IJ})$ or $\underbrace{X_I + X_1 + \mathcal{I}\mathcal{G}_{\mathcal{I}}^{IJ}\, C_J}_{\text{use direct CI methods}}$

$\downarrow$

merge $\mathcal{I}\mathcal{G}_{\mathcal{I}}^{IJ}$ to get H_{IJ}

Bender semi-direct CI and Shavitt unitary group are in this category.

B. Formula driven

1. For given I,J determine which $\mathcal{I}$ are needed

2. Extract these $\mathcal{I}$ from integral file

3. Form $H_{IJ} = \sum \mathcal{G}_{\mathcal{I}}^{IJ}$

C. Partial Formula Tape

1. For given I,J determine which $\mathcal{I}$ are needed and part of the $\mathcal{I}\mathcal{G}_{\mathcal{I}}^{IJ}$ formula

2. Sort partial formulas

3. Sort formulas on

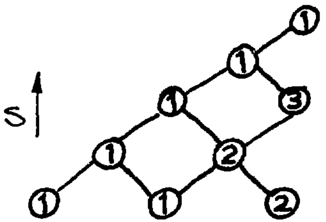
4. Form $\mathcal{I}\mathcal{G}_{\mathcal{I}}^{IJ}$ and store

5. Sort $\mathcal{I}\mathcal{G}_{\mathcal{I}}^{IJ}$ on I,J and combine to get H_{IJ}

D. Same as "C" except complete $\mathscr{G}^{IJ}$ is done in step 1.

SPIN-COUPLING

Number of different spin couplings



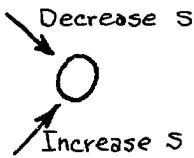
No. of different spin couplings

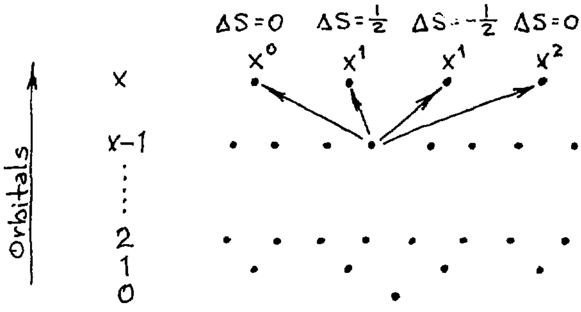
→ q = number of half-filled orbitals.

i.e., $S=0$ $q=2$ $(\alpha\beta-\beta\alpha)/\sqrt{2}$

$S=\frac{1}{2}$ $q=3$ $(\alpha\beta-\beta\alpha)\alpha/\sqrt{2}$

$(\alpha\beta\alpha + \beta\alpha\alpha - 2\alpha\alpha\beta)/\sqrt{2}$





Each path represents possible configuration

$$\Psi_{I\nu}^{SM} = \frac{1}{\sqrt{N!}} \mathcal{A}\{\phi_{i_1}\bar{\phi}_{i_1}\phi_{i_2}\bar{\phi}_{i_2} \cdots \phi_{i_p}\bar{\phi}_{i_p}\phi_{j_1} \cdots \phi_{j_q}\chi_\nu^{SM}\}$$

I labels $\{i_1 \ldots i_p;\ j_1 \ldots j_q\}$

χ_ν^{SM} is spin eigenfunction for q e^-.

$$\chi_\nu^{SM} = \sum_P A_{P,(S,\nu)}^{M,q} P(\alpha\alpha \ldots \alpha \quad \beta \ldots \beta)$$

$$\left\{\begin{array}{l} N_\alpha + N_\beta = q \\ N_\alpha - N_\beta = 2M \end{array}\right\} \rightarrow \text{P takes } \binom{q}{q+2M} \text{ values}$$

$$N_\alpha = q + 2M$$

Simplest χ has S=M and is usually used (E independent of M).

$$\Psi_I^{SM} = \sum_P A_{P,(S,\nu)}^{Mq} \mathcal{D}_{I,P}^{M} \qquad \text{or} \qquad A_{PS\nu}^{MqI}$$

$$\mathcal{D}_{I,P}^{M} = \frac{1}{\sqrt{N!}} \det\{\phi_{i_1}\bar{\phi}_{i_1} \cdots \phi_{i_p}\bar{\phi}_{i_p} \quad \phi_{j_1} \cdots \phi_{j_q}$$

$$\times \quad P(\alpha\alpha \ldots \beta \ldots)\}$$

$$\langle \Psi_{I\nu}^{SM} | H | \Psi_{I'\nu'}^{SM} \rangle = \sum_{P}^{\binom{q}{q+2M}} \sum_{P'}^{\binom{q'}{q'+2M}} A_P^q \quad A_{P'\nu'}^{q'} \langle \mathcal{D}_{IP}^{M} | H | \mathcal{D}_{I'P'}^{M} \rangle$$

Turn over rule

$$\exists \quad B_{p\nu}^q$$

$$<\Psi^{SM}_{I\nu}|H|\Psi^{SM}_{I'\nu'}> \;=\; \sum_{p}^{\lambda} \sum_{p'}^{(q')} B^{q}_{P\nu}\, A^{q'}_{P'\nu'} <\mathcal{O}^{M}_{Ip}|H|\mathcal{O}^{M}_{I'p'}>$$

where λ = number of spin couplings $<< \binom{q}{q+2M}$

EFFECT OF PERMUTING ORBITALS IN $\Psi_{I\nu}$

$\phi_{i_\lambda} \leftrightarrow \phi_{i_\mu}$	no effect
$\phi_{i_\lambda} \leftrightarrow \phi_{j_\mu}$	no effect
$\phi_{j_\lambda} \leftrightarrow \phi_{j_\mu}$	linear transform

i.e.,

$$\phi_1\phi_2\phi_3(\alpha\beta\alpha - \beta\alpha\alpha)/\sqrt{2} = \Psi_1$$

$$\phi_1\phi_2\phi_3(\alpha\beta\alpha + \beta\alpha\alpha + 2\alpha\alpha\beta)/\sqrt{6} = \Psi_2$$

Then suppose $\phi_1 \leftrightarrow \phi_3$

$$\phi_3\phi_2\phi_1(\alpha\beta\alpha - \beta\alpha\alpha)/\sqrt{2} = \Psi_1'$$

$$\phi_3\phi_2\phi_1(\alpha\beta\alpha + \beta\alpha\alpha - 2\alpha\alpha\beta)/\sqrt{2} = \Psi_2'$$

$$\Psi_1 = \frac{1}{2}\psi_1' + \frac{\sqrt{3}}{2}\psi_2$$

$$\Psi_2 = -\frac{\sqrt{3}}{2}\psi_1' + \frac{1}{2}\psi_2'$$

LINE UP PERMUTATION

$$i_1\bar{i}_1 i_2\bar{i}_2 \;\dots\; i_p\bar{i}_p j_1 \dots j_q$$

$$i_1'\bar{i}_1' i_2\bar{i}_2' \;\dots\; i_p' i_p' j_1' \dots j_q'$$

$\downarrow$

maximum juxtaposition

Puts I in "standard order" relative to I'; puts I' in order so mismatch is standardized.

$$\Psi_{I\nu} = \sum_{\bar{\nu}} Q_{\bar{\nu}\nu} \Psi^{I'}_{I\bar{\nu}}$$

$$<\Psi_{I\nu}|H|\Psi_{I'\nu'}> = \sum_{\bar{\nu}\bar{\nu}'} Q_{\bar{\nu}\nu} Q_{\bar{\nu}'\nu'} <\Psi^{I',I}_{I\bar{\nu}}|H|\Psi^{I'I}_{I'\bar{\nu}'}>$$

$$<\Psi^{I'I}_{I\bar{\nu}}|H|\Psi^{I'I}_{I'\bar{\nu}'}> = \sum_{p} \sum_{p'} B_p A_{p'} <\mathcal{D}^{I',I}_{Ip}|H|\mathcal{D}^{I',I}_{I'p'}> = \sum_{\mathcal{I}} \tilde{\mathcal{Y}}^{II'\bar{\nu}\bar{\nu}'}_{\mathcal{I}} \mathcal{I}$$

$$\mathcal{Y}^{II'\nu\nu'}_{\mathcal{I}} = \sum_{\bar{\nu}\bar{\nu}'} Q_{\bar{\nu}\nu} Q_{\bar{\nu}'\nu'} \tilde{\mathcal{Y}}^{II'\bar{\nu}\bar{\nu}'}_{\mathcal{I}}$$

Partial formula tape gives $\tilde{\mathcal{Y}}$; complete formula gives $\mathcal{Y}$.

STANDARDIZED CASES

A. Two orbitals differ (16 cases)

$i^2 \to k^2$, $\quad i^2 \to k\ell$, $\quad i^2k \to k^2\ell$, $\quad i^2k\ell \to k^2\ell^2$,

$i^2j \to ik^2$, $\quad i^2j \to ik\ell$, $\quad i^2jk \to ik^2\ell$, $\quad i^2jk\ell \to ik^2\ell^2$,

$i^2j^2 \to ijk^2$, $\quad i^2j^2 \to ijk\ell$, $\quad i^2j^2k \to ijk^2\ell$,

$i^2j^2k\ell \to ijk^2\ell^2$, $\quad ij \to k^2$, $\quad ij \to k\ell$, $\quad ijk \to k^2\ell$,

$ijk\ell \to k^2\ell^2$,

involves two $\mathcal{I}$ $[ik\|j\ell]$, $[i\ell\|jk]$. $\tilde{\mathcal{Y}}$ depends on case s,q.

B. One orbital differs (4 cases)

$i^2k \rightarrow ik^2, \qquad i \rightarrow k, \qquad i^2 \rightarrow ik, \qquad ik \rightarrow k^2$

involves h_{ik} $[ik\|pp]$, $[ip\|kp]$; all p common to I and I'.

C. I = I'

involves

$$\left\{ \sum_i n_i h_{ii} + \sum_i \sum_{j<i}^{n_j=2} n_i (2J_{ij} - K_{ij}) + \sum_i^{n_i=2} J_{ii} \right\}$$

and K_{ij} between open shells

<u>LARGE MATRIX EIGENVALUE</u> (Nesbit/Shavitt/Bender)

$$C^{(n)} = C^{(n-1)} + \lambda_i \hat{e}_i$$

$$\hat{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix}$$

2 × 2 CI

$$C^{(n-1)T} H C^{(n-1)} = E^{(n-1)}$$

$$\hat{e}_i^T H C^{(n-1)} = [H C^{(n-1)}]_i = \sum H_{ij} C_j^{(n-1)} = h_i$$

$$\hat{e}_i^T H \hat{e}_i = H_{ii}$$

$$\begin{pmatrix} E^{(n-1)} & h_i \\ h_i & H_{ii} \end{pmatrix} \begin{pmatrix} 1 \\ \lambda_i \end{pmatrix} = E^{(n)} \begin{pmatrix} 1 \\ \lambda_i \end{pmatrix}$$

$$\left(E^{(n)} - E^{(n-1)}\right)\left(E^{(n)} - H_{ii}\right) = h_i^2$$

$$\left(E^{(n)} - E^{(n-1)}\right) \simeq - \frac{h_i^2}{H_{ii} - E^{(n-1)}} \qquad \text{(Nesbit)}$$

$$E^{(n)} - E^{(n-1)} = \frac{-h_i^2}{\frac{1}{2}\left[\left(H_{ii} - E^{(n-1)}\right) + \sqrt{\left(H_{ii} - E^{(n-1)}\right)^2 + 4h_i^2}\right.}$$

$$\lambda_i = \frac{h_i}{\tilde{E}^{(n)} - H_{ii}}$$

$$E^{(n)} = \langle H \rangle = \frac{E^{(n-1)} + 2h_i\lambda_i + H_{ii}\lambda_i^2}{1 + \lambda_i^2}$$

Difficult for excited states

$$\underset{\sim}{C}^{(n)} = \sum_{i=1}^{k} \alpha_i \underset{\sim}{b}_i$$

$$\tilde{H}_{ij} = b_i^{+} H \underset{\sim}{b}_j \,, \qquad \underset{\sim}{b}_i^{+} \underset{\sim}{b}_j = \delta_{ij}$$

$$\underset{\sim}{\tilde{H}}\, \underset{\sim}{\alpha} = E_k \underset{\sim}{\alpha} \,, \qquad k^{\underset{\sim}{\alpha}} \text{ eigenvalue}$$

$$q^{(n)} = (H - E_k)\, C^{(n)}$$

$$\xi_p^{(n)} = (E_k - H_{pp})^{-1}\, q_p^{(n)}$$

First order correction to $C^{(n)}$. Schmidt orthogonalize ξ to $\{b\}$ and normalize $\longrightarrow b_{k+1}$. Iterate until $C^{(n)}$ converges.

PERTURBATION CORRECTIONS AND APPROXIMATIONS

Quadruple Excitations

$$\Delta E_Q \;\cong\; (1 - C_o^2)\; \Delta E_{SD}$$

CEPA: Shift diagonal elements to account for quad.

Pair correlation ij →

Shift by $\delta H \cong \Delta E$ all other peaks.

Segal/Davidson/Shavitt B_k: Neglect most off-diagonal elements

$$\underset{\sim}{H} \;\cong\; \left(\begin{array}{c|c} H_o & h^T \\ \hline h & D \end{array}\right) \qquad\qquad C \;\cong\; \left(\begin{array}{c} C_o \\ \hline x \end{array}\right)$$

$$H_oC_o + h^Tx = EC_o$$

$$hC_o + D\,x \cong E\,x$$

$$x \cong (E - D)^{-1}\, hC_o$$

$$[H_o + h^T(E - D)^{-1}\, h]C_o = EC_o$$

nonlinear in E.

$$(F\text{-}D)^{-1} \equiv (E_o\text{-}D)^{-1} - (E\text{-}E_o)(E_o\text{-}D)^{-1}(E\text{-}D)^{-1}$$

$$[H_o\text{-}E_o + h^T(E_o\text{-}D)^{-1}h]C_o = (E\text{-}E_o)[\mathbf{1} + h^T(E_o\text{-}D)^{-1}(E\text{-}D)^{-1}h]C_o$$

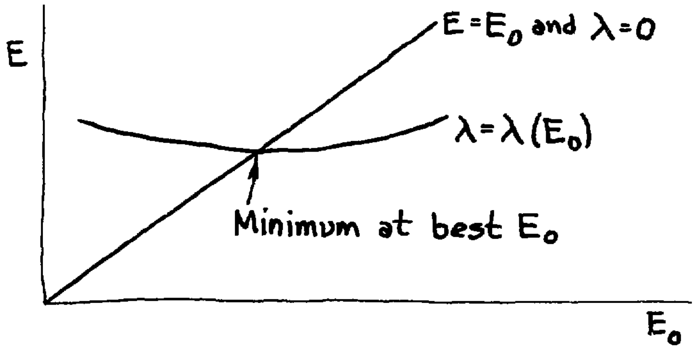
Solve as simultaneous equations

$$\tilde{H} = H_o - E_o + \underline{h}^T (E_o - D)^{-1} \underline{h}$$

$$\tilde{S} = 1 + \underline{h}^T (E_o - D)^{-2} \underline{h}$$

$$\tilde{H} C_o = \lambda \tilde{S} C_o$$

where $\lambda = 0$ and $E = E_o + \lambda$



since $\partial E / \partial E_o = 0$ at $E = E_o$, small errors in E_o have little effect on E.

GEOMETRY OPTIMIZATION OF LARGE SYSTEMS

Lecture 14

by

Michael C. Zerner

Guelph-Waterloo Centre for Graduate Work in Chemistry
University of Guelph
Guelph, Ontario, Canada

INTRODUCTION

Perhaps one of the most successful applications of molecular quantum mechanics has been the reproduction and prediction of molecular conformation. In many cases bond lengths are reproduced to ±0.02 Å and bond angles to ±5° with a variety of simple molecular orbital models, or with minimum basis set *ab initio* calculations.[1,2] Larger basis sets, especially those of double ζ plus polarization type and the inclusion of electron correlation are now producing geometries which challenge crystallography for accuracy. The optimist, armed with the growing success of conformational calculations, might even choose the calculated results on isolated molecules over the experimental results obtained in condensed media, as the former may be more appropriate for the chemistry he is investigating.

In addition to yielding information about global minima of the potential energy surface, quantum mechanical calculations yield information on local minima, which may or may not be observable directly, but which might be involved in reaction pathways. Similarly information can be obtained about transition states and energy barriers that would be difficult or impossible to obtain in other ways.

The gleaning of all this information from a potential energy surface is difficult. Considering N atoms there are 3N-6 (or 3N-5) degrees of freedom that should be plotted against the energy. For detailed statistical calculations one may have to live with this "3N" problem and visit all regions of the surface thermally accessible. These notes, however, are concerned with determining only a small part of the potential energy surface: those points that either correspond to minima, and thus stable or metastable conformations, and points that correspond to transition states.

GENERAL CONSIDERATIONS

The energy E of a molecular system obtained under the Born-Oppenheimer approximation is a parametric function of the coordinates $X = (X_1, X_2, \ldots X_{3N})$ assumed for the calculation. We wish to move from $E(\underline{X})$ to $E(\underline{X}^1)$, where $\underline{q} = (\underline{X}^1 - \underline{X})$. This may be summarized as a Taylor expansion about $\underline{X}$ as

$$E(\underline{X}^1) = E(\underline{X}) + \underline{f}(\underline{X})\underline{q}^+ + \tfrac{1}{2}\,\underline{q}\underline{H}(\underline{X})\underline{q}^+ + \ldots \qquad (1)$$

with

$$f_i = \frac{\partial E(\underline{X})}{\partial X_i}$$

and

$$H_{ij} = \frac{\partial E(\underline{X})}{\partial X_i \partial X_j}$$

the gradient $\underline{f}$ and Hessian $\underline{H}$ matrices, respectively. Although conceptually the Taylor series is infinite, about extrema we might expect a quadratic form to be adequate; i.e., for $\underline{X} = \underline{X}_e$, where $\underline{X}_e$ designates a stationary point and by definition is characterized by $\underline{f}(\underline{X}_e) = \underline{0}$,

$$E(\underline{X}) = E(\underline{X}_e) + \tfrac{1}{2}\,\underline{q}\underline{H}(\underline{X}_e)\underline{q}^+ + \ldots$$

In a similar fashion

$$\underline{f}(\underline{X}^1) = \underline{f}(\underline{X}) + \underline{q}\underline{H}(\underline{x}) + \ldots \tag{2}$$

For the point $\underline{X}^1 = \underline{X}_e$,

$$\underline{f}(\underline{X}) = -\underline{q}\underline{H}(\underline{X}) \tag{3}$$

The solution of Eq. (3) is the starting point of the most efficient procedures used to find extrema in functions of several variables where the functional form of $E(\underline{X})$ is not explicit in $\underline{X}$. If $\underline{H}$ is non-singular, then

$$\underline{q} = -\underline{f}(\underline{X})\ \underline{H}^{-1}(\underline{X}) \tag{4}$$

which allows the solution for $\underline{X}_e$ from any point $\underline{X}$ near enough so that the energy function is nearly quadratic. Similarly, an estimate of $E(\underline{X}_e)$ is obtained from

$$\begin{aligned} E(\underline{X}_e) &= E(\underline{X}) - \tfrac{1}{2}\underline{f}(\underline{X})\ H^{-1}(X)\ f^+(\underline{X}) \\ &= E(\underline{X}_e) - \tfrac{1}{2}\underline{q}H(\underline{X})\ \underline{q}^+ \end{aligned} \tag{5}$$

For the specific problem of uncovering extrema on the potential energy surface there are several pathological considerations. The first of these is that $H^{-1}(X)$ will not exist unless the rotations and translations which represent zero eigenvalues of $\underline{H}$ have been factored. This may be accomplished via the $\underline{B}$ matrix of Wilson and Eliashevich[3]

$$\underset{\sim}{Y} = \underset{\sim}{X}\,\underset{\sim}{B} \tag{6}$$

where $\underset{\sim}{X}$ has 3N entries and $\underset{\sim}{B}$ is $3N \times 3N\text{-}6$. Since work, W, is independent of the choice of coordinate systems, the six (or five) zero forces can be separated by

$$W = \underset{\sim}{f}\,\underset{\sim}{q}^{+} = \underset{\sim}{f}_y\,\underset{\sim}{q}_y^{+} = \underset{\sim}{f}_y\,\underset{\sim}{B}^{+}\,\underset{\sim}{q}^{+}$$

or

$$\underset{\sim}{f} = \underset{\sim}{f}_y\,\underset{\sim}{B}^{+} \qquad \underset{\sim}{f}_y = \underset{\sim}{f}(\underset{\sim}{B}^{+})^{-1} \tag{7}$$

where $(B^{+})^{-1}$ is defined from

$$\underset{\sim}{B}^{+}(\underset{\sim}{B}^{+})^{-1} = \underset{\sim}{1}$$

In general, $(B^{+})^{-1}$ can be given by

$$(\underset{\sim}{B}^{+})^{-1} = \underset{\sim}{m}\,\underset{\sim}{B}(\underset{\sim}{B}^{+}\underset{\sim}{m}\,\underset{\sim}{B})^{-1}$$

where $\underset{\sim}{m}$ is an arbitrary $3N \times 3N$ matrix, usually taken as a diagonal matrix containing the reciprocal of each atomic mass three times in the appropriate positions.[4] It may also be chosen as the unit matrix with six (or five) zero entries chosen to prevent translation and rotation. A simple such choice of this type is to place atom #1 at the origin, atom #2 on the z-axis and atom #3 in the xz-plane. Then the six (or five) coordinates removed are $X_1=Y_1=Z_1=0$, $X_2=Y_2=0$, and $Y_3=0$. If $Y_3=0$ implies $X_3=0$ for any choice of third atoms, then the molecule is linear and only five degrees of freedom are chosen.

In addi ion to the consideration that must be given to the inverse of $\underset{\sim}{H}$, and to which we shall return, we must recall that neither $\underset{\sim}{f}$ nor $\underset{\sim}{H}$ are generally calculated in quantum chemical computations. For this we must consider the energy E and how it is obtained. For Hartree-Fock calculations E is dependent explicitly on the occupied molecular orbital coefficients $\underset{\sim}{C}$ and on $\underset{\sim}{X}$. Its derivation is then given by

$$\frac{dE}{dX_A} = \frac{\partial E}{\partial X_A} + \sum_{i,a} \frac{\partial E}{\partial C_{ia}} \frac{\partial C_{ia}}{\partial X_a} \tag{8}$$

Since $\partial E/\partial C_{ia} = 0$ is the condition for the Hartree-Fock solutions,

$$\frac{dE}{dX_\alpha} = \frac{\partial E}{\partial X_\alpha} \tag{9}$$

This realization allows one to ignore to first order the change in molecular orbital coefficients with respect to geometry changes. Given for a closed shell system that

$$E = \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu} + \sum_{\mu\nu\sigma\lambda} P_{\mu\nu} P_{\sigma\lambda} \langle\mu\lambda\|\nu\sigma\rangle + V_{NN} \tag{10}$$

where

$$V_{NN} = \sum_{A<B} \frac{Z_A Z_B}{R_{AB}}$$

$$P_{\mu\nu} = \sum_{a}^{\text{m.o.}} C_{\mu a} C_{\nu a} n_a$$

$$\begin{aligned}\langle\mu\lambda\|\nu\sigma\rangle &= \Big(\mu(1)\nu(1)\,|\,\lambda(2)\sigma(2)\Big) \\ &\quad - \tfrac{1}{2}\Big(\mu(1)\sigma(1)\,|\,\lambda(2)\mu(2)\Big)\end{aligned}$$

where $\underline{h}$ is the one-electron matrix, Z_A the atomic number of atom A, and $\underline{P}$ is the first order density. Differentiating Eq. (10) yields

$$\frac{\partial E}{\partial X_A} = \sum_{\mu\nu} \frac{\partial h_{\mu\nu}}{\partial X_A} + \sum_{\mu\nu\sigma\lambda} P_{\mu\nu} P_{\sigma\lambda} \frac{\partial \langle\mu\lambda\|\nu\sigma\rangle}{\partial X_A} + \frac{\partial V_{NN}}{\partial X_A}$$

$$+ \sum \frac{\partial P_{\mu\nu}}{\partial X_A} h_{\mu\nu} + 2 \sum_{\mu\nu\sigma\lambda} \frac{\partial P_{\mu\nu}}{\partial X_A} P_{\sigma\lambda} \langle\mu\lambda\|\nu\sigma\rangle \qquad (11)$$

Equation (11) suggests that derivatives of the m.o. coefficients are required, whereas Eq. (9) does not! Expanding the last two terms of Eq. (11) gives

$$2 \sum_{\mu,\nu} \sum_{a}^{\text{m.o.}} \frac{\partial C_{\mu a}}{\partial X_A} h_{\mu\nu} C_{\nu a} n_a + 2 \sum_{\mu,\nu,\sigma,\lambda} \sum_{a}^{\text{m.o.}} \frac{\partial C_{\mu a}}{\partial X_A} P_{\sigma\lambda} \langle\mu\sigma\|\nu\lambda\rangle C_{\nu a} n_a$$

$$= 2 \sum_{\mu,\nu} \sum_{a}^{\text{m.o.}} \frac{\partial C_{\mu a}}{\partial X_A} \left\{ h_{\mu\nu} + \sum_{\sigma,\lambda} P_{\sigma\lambda} \langle\mu\sigma\|\nu\lambda\rangle \right\} C_{\nu a}\, n_a$$

$$= 2 \sum_{\mu,\nu} \sum_{a}^{\text{m.o.}} \frac{\partial C_{\mu a}}{\partial X_A} \{F_{\mu\nu}\} C_{\nu a}\, n_a$$

$$= 2 \sum_{a}^{\text{m.o.}} \sum_{\mu\nu} \frac{\partial C_{\mu a}}{\partial X_a} \varepsilon_a S_{\mu\nu} C_{\nu a}\, n_a$$

Recalling that the orthonormality condition of the molecular orbitals are

$$\sum_{\mu,\nu} C_{\mu a}\, S_{\mu\nu}\, C_{\nu b} \;=\; \delta_{ab}$$

and differentiating the above yields

$$2\sum_{\mu,\nu} \frac{\partial C_{\mu a}}{\partial X_A}\, S_{\mu\nu}\, C_{\nu a} \;=\; -2\sum_{\mu,\nu} C_{\mu a}\, C_{\nu a}\, \frac{\partial S_{\mu\nu}}{\partial X_A}$$

which has eliminated the derivatives of the coefficients. Combining these expressions results in

$$\frac{\partial E}{\partial X_A} = \sum_{\mu\nu} P_{\mu\nu}\, \frac{\partial h_{\mu\nu}}{\partial X_A} + \sum_{\mu\nu\sigma\lambda} P_{\mu\nu} P_{\sigma\lambda}\, \frac{\partial \langle\mu\lambda\|\nu\sigma\rangle}{\partial X_A}$$

$$- \sum_{\mu,\nu} P'_{\mu,\nu}\, \frac{\partial S_{\mu\nu}}{\partial X_A} + \frac{\partial V_{NN}}{\partial X_A} \qquad (12)$$

where

$$P'_{\mu\nu} \;=\; \sum_{a}^{\text{m.o.'s}} \varepsilon_a\, C_{\mu a}\, C_{\nu a}\, n_a \quad .$$

The relative simplicity of Eq. (12), with no derivatives of $\underset{\sim}{P}$ appearing should not be confused with the Hellman-Feynman theorem.[5] Given that

$$E \;=\; \langle\psi|H|\psi\rangle$$

with $\langle\psi|\psi\rangle = 1$, then

$$\frac{dE}{dX_A} \;=\; \langle\frac{\partial\psi}{\partial X_A}|H|\psi\rangle + \langle\psi|H|\frac{\partial\psi}{\partial X_A}\rangle + \langle\psi|\frac{\partial H}{\partial X_A}|\psi\rangle \qquad (13)$$

The Hellman-Feynman condition then is that

$$< \frac{\partial\psi}{\partial X_A} |H|\psi> + <\psi|H| \frac{\partial\psi}{\partial X_A} > \quad = \quad 0 \tag{14}$$

which only holds for exact solutions, or certain classes of trial functions.[6] Under the constraints of Eq. (14), Eq. (13) is simply

$$\frac{dE}{dX_A} \quad = \quad <\psi| \frac{\partial H}{\partial X_A} |\psi> \quad . \tag{15}$$

Equation (15) is the expectation valve of a simple one-electron operator plus the derivative of the nuclear repulsion term. Equation (11), however, does not depend on Eq. (14). The integrals involved in $\partial \underline{h}/\partial X_A$ and $\partial(\mu\nu|\sigma\lambda)/\partial X_A$, discussed later, involve the wavefunction through "atomic orbital following," i.e., $\partial \chi_a/\partial X_A$, with χ_a an atomic orbital on center A, and are far more complicated than those of Eq. (15). In practice, the forces evaluated through Eq. (15) can be large even when they are calculated to be zero under Eq. (13) and thus represent an extrema of the energy function. Nevertheless, the simplicity of Eq. (15) is appealing, and one wonders if the increased inconvenience of insuring Eq. (14) is not repaid in utilizing Eq. (15) when the goal is geometry optimization.[7]

For a configuration interacted (CI) wavefunction over determinants ψ_j,

$$\psi_T \quad = \quad \sum_j d_j \ \psi_j$$

one obtains for the energy derivatives

$$\frac{dE}{dX_A} = \frac{\partial E}{\partial X_A} + \sum_{ib} \frac{\partial E}{\partial C_{ib}} \frac{\partial C_{ib}}{\partial X_A} + \sum_{j} \frac{\partial E}{\partial d_j} \frac{\partial d_j}{\partial X_A} \tag{16}$$

where now the first sum is over all molecular orbital coefficients. In this case, $dE/dX_A = \partial E/\partial X_\alpha$ only for a multi-configuration self-consistent field (MCSCF) function. For the general Hartree-Fock plus CI wavefunction, $\partial E/\partial d_i = 0$ and

$$\frac{dE}{dX_\alpha} = \frac{\partial E}{\partial X_\alpha} + \sum_{i,a} \frac{\partial E}{\partial C_{ia}} \frac{\partial C_{ia}}{\partial X_\alpha} \tag{17}$$

The evaluation of $\partial C_{ia}/\partial X_\alpha$ is complicated, but can be approached through perturbation theory.[8] The contribution to the forces of the second term might be expected to be small for a large CI, as the dependence of the energy on $\underset{\sim}{C}$ is downgraded, or for a system without a great deal of bond polarity, or for a system in which the molecular orbitals are determined by symmetry. Under such situations an initial search can be made of the surface using Eq. (9), but for accurate results reliance on this approximation is not satisfactory.

Second derivatives of the Hartree-Fock energy can be obtained directly from Eq. (12):

$$\frac{\partial^2 E}{\partial X_A \partial X_B} = \sum_{\mu,\nu} P_{\mu\nu} \frac{\partial^2 h_{\mu\nu}}{\partial X_A \partial X_B} + \sum_{\mu\nu\sigma\lambda} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial^2 \langle\mu\lambda\|\nu\sigma\rangle}{\partial \quad \partial}$$

$$- \sum_{\mu,\nu} P'_{\mu\nu} \frac{\partial^2 S_{\mu\nu}}{\partial X_A \partial X_B} + \frac{\partial^2 V_{NN}}{\partial X_A \partial X_B} + \sum_{\mu,\nu} \frac{\partial P_{\mu\nu}}{\partial X_B} \frac{\partial h_{\mu\nu}}{\partial X_A}$$

$$+ 2\sum_{\mu,\nu,\sigma,\lambda} P_{\mu\nu} \frac{\partial P_{\lambda\sigma}}{\partial X_B} \frac{\partial(\mu\nu|\sigma\lambda)}{\partial X_A} - \sum \frac{\partial P'_{\mu\nu}}{\partial X_B} \frac{\partial S_{\mu\nu}}{\partial X_A}$$

The last three terms of this expression involve the derivatives of the molecular orbital coefficients and cannot easily be avoided.

OPTIMIZATION TECHNIQUES

General Considerations

There is a rather large literature on numerical methods for finding stationary points as a function of many variables.[9,10] For the purposes of these notes they may be classified as follows:

a) methods without gradients

b) methods with numerical gradients and second derivatives

c) methods with analytical gradients and numerical second derivatives

d) methods with analytical gradients and analytical second derivatives.

All of these methods relate to the Taylor expansions of the function E and its derivatives $\underline{f}$ as given in the previous section. In practice, they can be applied as "estimate" techniques, or as "iterative" techniques.

Type (d) methods might be preferred as utilizing the maximum amount of information at a given point, but assumes that the analytic first and second derivatives can be obtained at the same time, and with the same ease, as the energy E. It is clear, however, that insofar as our initial estimate of the geometry at an extrema is within the quadratic region of the valence bond force field (the $\underline{y}$ coordinates of Eq. (6)), a single application of Eqs. (4) and (5) give a set $\underline{y}_e$ and the energy $E(\underline{y}_e)$. Such a single application of Eq. (4) we shall call an *estimate*. If we are not within the quadratic region of the potential, the estimate may not be very accurate, and it may be desirable to *iterate*; that is, having determined a new set, $\underline{y}^1$, from the initial guess $\underline{y}_0$, solve the equations of the previous section for $\underline{y}^2$. This requires $\underline{f}(\underline{y}^1)$ and $\underline{H}^{-1}(\underline{y}^1)$. This procedure might then be repeated until $E_n - E_{n-1}$ is below a given threshold or $\sigma = \underline{f}(\underline{y}^n)\ \underline{f}^+(\underline{y}^n)$ is below a given threshold, or both.

In practice, type (d) algorithms are not used because of the difficulty that arises in analytically obtaining the required derivatives. In general, the derivative of an orbital with respect to a nuclear coordinate gives rise to several new orbitals (see below), at least one of which is of greater ℓ quantum number than the orbital itself. The second derivatives will involve even more terms, with atomic orbitals of value $\ell+2$. Assuming an SCF calculation requires s integrals, ~5s or more integrals are required for the derivatives $\underline{f}$, and ~25s for the second derivatives. Since $s \sim n^4$ for *ab initio* methods, where n is the number of basis atomic orbitals, and the SCF step proceeds as n^3, integral evaluation is already the time-consuming step. It might thus be possible to perform many SCF calculations in the same time required to evaluate $\underline{f}$ analytically. On the other hand, most semi-empirical methods have $\sim n^2$ integrals required in the formation of the Fock matrix. This time-consuming step is the solution of the secular equation, and the evaluation of analytic first derivatives are quickly accomplished. The most efficient methods used today are of Type (c), but certainly attempts to utilize Type (d) are in order for methods in which integral evaluation is not time-consuming.

Some Algorithms

The simplest of the methods are of Type (a). The simplest of these are the so-called axial iteration or univariant techniques. One chooses a set of internal coordinates and minimizes the potential energy with respect to each coordinate in turn. After completing the $3N-6 = m$ independent searches, one returns and repeats the procedure until the change in coordinates is below a given threshold.

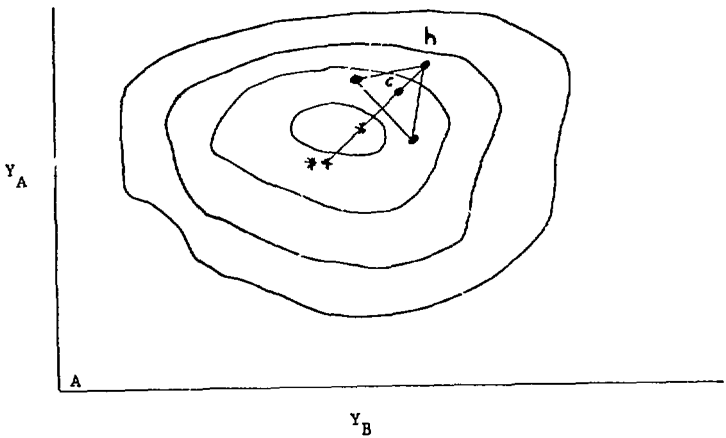
One successful such procedure is to step alone each coordinate y_i by α_i. If $E(\underline{Y} + \alpha_i \underline{e}_i) < E(\underline{Y})$, where $\underline{e}_i$ is the unit vector along "i", repeat the step until $E(\underline{Y} = r\alpha_i \underline{e}_i) >$ $E(\underline{Y} + (r-1)\alpha_i \underline{e}_i)$, r an integer. The new coordinates are $\underline{Y} = \underline{Y} + (r-1)\alpha_i \underline{e}_i$. If $E(Y + \alpha_i \underline{e}_i) > E(\underline{Y})$, step the other direction until $E(Y - r\alpha_i \underline{e}_i) > E(\underline{Y} - (r-1)\alpha_i \underline{e}_i)$. Again the new coordinates are $Y = Y - (r-1)\alpha_i \underline{e}_i$. If $E(\underline{Y})$ is of lower energy than both $E(Y + \alpha_i \underline{e}_i)$ and $E(\underline{Y} - \alpha_i \underline{e}_i)$, then a quadratic is fit through the three points $(y_i, E(y_i))$ and the minimum value of the quadratic found $(y_i^*, E(y_i^*))$. The coordinates $\underline{Y}$ are updated and α_i is set to $\alpha_i/4$. This procedure is repeated for all i, and then iterated until all α_i are below a specified threshold.

The most effective of the Type (a) algorithms seem to be of the Simplex type.[11,12,13] The method given below is that of Nelder and Mead.[13] Figure 1 is a schematic attempt to follow this method for two variables.

Consider m variables. $\underline{X}_0, \underline{X}_1, \ldots \underline{X}_m$ are the m+1 independent points in this $\underline{m}$ dimensional space that defines the "simplex." E_i designates the value of the energy $E(\underline{X}_i)$. Let E_h be the highest value of $\{E_i\}$, and E the lowest. Let $\overline{\underline{X}}$ be the centroid of the points $\{\overline{X}_i\}_{i \neq h}$ and $[\underline{X}_i\ \underline{X}_j]$ the distance between $\underline{X}_i$ and $\underline{X}_j$.

$$[\underline{X}_i\ \underline{X}_j]^2 = \sum_{a=1}^{m} [X_i(a) - X_j(a)]^2$$

The reflection of $\underline{X}_h$ is denoted $\underline{X}^*$ and its coordinates given by



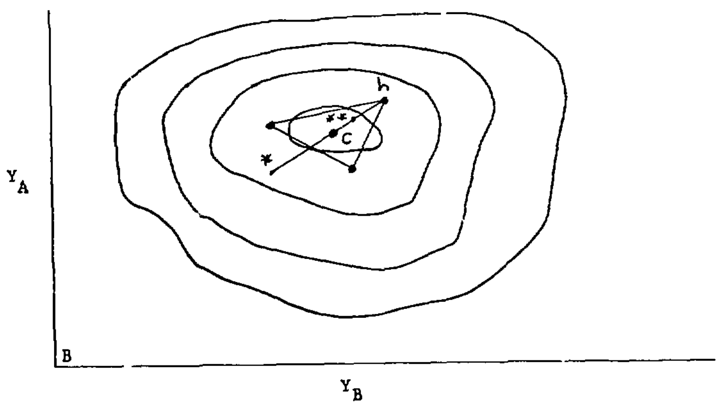


Fig. 1. The Simplex Method, where h designates E_H, C designates the centroid of points: A, a successful reflextion *, but failed expansion **; B, a failed reflection *, but successful contraction.

$$X^* = (1 + \alpha)\overline{\underline{X}} - \alpha \underline{X}_h$$

$$\alpha = \frac{[\underline{X}^* \overline{\underline{X}}]}{[\underline{X}_h \overline{\underline{X}}]} \tag{18}$$

where α is called the reflection constant and is positive. $\underline{X}^*$ is thus on a line joining $\underline{X}_h$ and $\overline{\underline{X}}$, but reflected to the far side of $\overline{\underline{X}}$ from $\underline{X}_h$. Three possibilities ensue: If $E_\ell < E^* < E_h$, then $\underline{X}_h$ is replaced by $\underline{X}^*$ and one starts again with a new simplex, reflecting the new $\underline{X}_h$, etc.

If $E^* < E_\ell$, if the reflection has produced a new minimum, then $\underline{X}^*$ is *expanded* to $\underline{X}^{**}$ by

$$X^{**} = \nu X^* + (1 - \nu)\overline{\underline{X}} \tag{19}$$

where

$$\nu = \frac{[\underline{X}^{**} \overline{\underline{X}}]}{[\underline{X}^* \overline{\underline{X}}]} > 1$$

where ν is the *expansion* coefficient. If $E^{**} < E_\ell$, $\underline{X}_h$ is replaced by $\underline{X}^{**}$ and the procedure is restated. If $E^{**} > E_\ell$, then the expansion has failed and $\underline{X}_h$ is replaced by $\underline{X}^*$ before restarting.

Finally, if $E^* > E_i$ for all $i \neq h$, that is, replacing $\underline{X}_h$ with $\underline{X}^*$, leaves $\underline{E}^*$ the new maximum; then a new $\overline{X}_h$ is defined which produces the minimum of $E(\overline{\underline{X}}_h)$ and $E(\underline{X}^*)$ and a new *contraction* is examined

$$\underline{X}^{**} = \beta \underline{X}_h + (1 - \beta)\overline{\underline{X}} \tag{20}$$

$$\beta = \frac{[\underline{X}^{**}\overline{\underline{X}}]}{[\underline{X}_n\ \overline{\underline{X}}]}$$

The *contraction* coefficient β lies between 0 and 1. $\underline{X}^{**}$ replaces $\overline{\underline{X}}_h$ unless $E^{**} > \min(\underline{X}_h, \underline{X}^{*})$. In the latter rather rare case, all $\underline{X}_i$ are replaced by $(\underline{X}_i + \underline{X}_\ell)/2$ and the process restarted.

From an analysis of analytic functions Nelder and Mead suggest the values $\alpha = 1$, $\beta = \frac{1}{2}$, $\nu = 2$. For the mathematical implications of this strategy one is referred to the original literature.[13]

Applications of the simplex method to molecular orbital calculations have proven reasonably successful. The MINDO/2 method, for example, was parameterized by such a geometry optimization procedure.[14] A strong advantage of the simplex method over axial or invariant methods appears when the number of variables becomes large, and coupling between these variables are large.

The most successful methods that use gradients, either numerical or analytic, that I have examined seem to rest on the Murtagh Sargent[15] variant of the Davidon[16] Fletcher-Powell[17] method. This procedure is appropriate for Type (c) or Type (d) algorithms, and proceeds as we have implemented it, as follows:[18]

A sequence $\underline{S}_n = \underline{S}(\underline{X}_n)$ is defined that will approach $H^{-1}(\underline{X}_e)$ for sufficiently large n. The starting point is Eq. (4).

$$q_K = \underline{X}_K - \underline{X}_{K-1} = -c_{K-1}\,\underline{f}_{K-1}\,\underline{S}_{K-1} \qquad (21)$$

Step 1: Set $\alpha_0 = \frac{1}{2}$ and $\underline{S}_0 = 1$. Use Eq. (21) to obtain a new set of coordinates $\underline{X}_1$. If $E_1 > E_0$, repeat this step with $\alpha_0 = \frac{1}{2}\alpha_0$ until $E_1 < E_0$. This is equivalent to the method of steepest descent with "half-steps."

Step 2: Form

$$\begin{aligned}\underline{Z}_K &= -\alpha_{K-1}\,\underline{f}_{K-1}\,\underline{S}_{K-1} - (\underline{f}_K - \underline{f}_{K-1})\underline{S}_{K-1} \\ &= [\underline{f}_K + \underline{f}_{K-1}(\alpha_{K-1} - 1)]\underline{S}_{K-1} \\ C_K &= \underline{Z}_K(\underline{f}_K - \underline{f}_{K-1}) \\ CT_K &= \underline{Z}_K\,\underline{Z}_K^{+}\end{aligned} \tag{22}$$

If $|C_K| < 10^{-5}\,CT_K$ or $\underline{Z}_K\underline{f}_{K-1}^{+}/C_K > 10^{-5}$, S_K is reset to $\underline{1}$, α_K reset to $\frac{1}{2}$. These tests insure the stability of $\underline{S}_K$; that is, the $\underline{S}_K$ remains positive definite after update. If the rotational and translational degrees of freedom have been removed via $\underline{B}$ of Eq. (6), then these tests might fail because of the numerical updating procedure. If these degrees of freedom have not been removed, then $\underline{H}^{-1}(X_e)$ is indeed singular and eventually the procedure will recognize this. In either case, with reasonable starting geometries, $\underline{S}_K$ is seldom reset to $\underline{1}$, a fortunate finding, for this reset would mean the loss of all information about the curvature of the surface built up from previous cycles. If these two tests are passed, then $\underline{S}_{K-1}$ is updated by

$$\underline{S}_K = \underline{S}_{K-1} + \underline{Z}_K^{+}\,\underline{Z}_K/C_K \tag{23}$$

and α_K is set to unity.

Step 3: Equation (21) is solved to find a new set of $\underline{X}_K$. E_K and $\underline{f}_K$ are calculated. If $E_K < E_{K-1}$, one reports back to step 2 until $\sigma/m = \underline{f}_K \underline{f}_K^{+}/m < 10^{-3}$ a.u., at which point most bond lengths are converged to ±0.01 a.u. and bond angles to ±0.5°. If $E_K > E_{K-1}$, $\alpha_{K-1} = \frac{1}{2}\alpha_K$ (a "backstep"), and step 2 is repeated without updating K.

An important feature of the Murtagh Sargent procedure is that a stationary value of E is obtained at the latest in m+1 steps *even if* H *is singular.*[15,19] In practice , far fewer steps than m+1 are required if reasonable guesses on starting geometries are available.

FIRST DERIVATIVES OF THE ENERGY

Algorithms that utilize the first derivatives in searching for extrema are, in general, more effective than those that do not. The question then concerns the difficulty in obtaining these gradients.

Numerical Methods

The derivatives $\underline{f}$ can always be obtained by central difference

$$f_i \equiv \frac{\partial E}{\partial X_i} \cong \frac{E(\underline{X} + \alpha_i \underline{e}_i) - E(\underline{X} - \alpha_i \underline{e}_i)}{2\alpha_i} \qquad (24)$$

where α_i defines the step size and $\underline{e}_i$ is a unit vector in the

i^{th} direction. The step size must be chosen large enough such that the difference $E(\underline{X} + \alpha_i \underline{e}_i) - E(\underline{X} - \alpha_i \underline{e}_i)$ is numerically stable, and small enough such that this finite difference equation approaches the derivative. In practice $\alpha_1 = 0.05$ Å and 1° seem satisfactory,[20,21] although smaller values have been suggested.[22] Assuming m degrees of freedom ($\underline{X}$ in a $1 \times m$ row vector), $2m+1$ calculations must be performed to obtain $\underline{f}$ for each cycle of the Type (b) or Type (c) algorithms.

A more accurate estimate can be made of these derivatives from the Newton-Stirling formula[23,24]

$$\alpha f_i \equiv \alpha \left(\frac{\partial E(\underline{X})}{\partial X_i} \right)_{\underline{X}=\underline{X}_0} = \mu\delta E_0 - \frac{\mu\delta E_0^3}{3!} + \ldots$$

$$+ \frac{(-1)^n (n!)^2}{(2n+1)!} \mu\delta E_0^{2n+1} \tag{25}$$

$$\delta E_r^{2n} = \delta E_{r+1}^{2n-2} - 2\delta E_r^{2n-2} + \delta E_{r-1}^{2n-2}$$

$$\mu\delta E_r^{2n-1} = \tfrac{1}{2}\left(\delta E_{r+1}^{2n-2} - \delta E_{r-1}^{2n-2}\right)$$

$$\delta E_r^0 = E(\underline{X} + r\alpha \underline{e}_K) , \qquad r = \ldots -2, -1, 0, 1, 2, \ldots$$

The leading term in Eq. (25) is the central difference formula of Eq. (24). The extra expense involved in going beyond the first term in Eq. (25) hardly seems worthwhile in obtaining the elements of $\underline{f}$ if a reasonable starting geometry has been guessed.

Analytic Derivatives

From Eq. (11), for a closed shell molecule, we obtain

$$\frac{\partial E}{\partial X_i} = \sum_{\mu,\nu} \left(\mu \left| \frac{\partial V_{NE}}{\partial X_i} \right| \nu\right) P_{\mu\nu}$$

$$+ \sum_{\mu,\nu} \left[< \frac{\partial \mu}{\partial X_i} | V_{NE} | \nu > + < \mu | V_{NE} | \frac{\partial \nu}{\partial X_i} > \right] P_{\mu\nu}$$

$$+ \sum_{\mu,\nu} \left[< \frac{\partial \mu}{\partial X_i} | T | \nu > + < \mu | T | \frac{\partial \nu}{\partial X_i} > \right] P_{\mu\nu}$$

$$+ \sum_{\mu\nu\sigma\lambda} \frac{\partial}{\partial X_i} <\mu\lambda \| \nu\sigma> P_{\mu\nu} P_{\sigma\lambda}$$

$$- \sum_{\mu\ \nu} \left[< \frac{\partial \mu}{\partial X_i} | \nu > + < \mu | \frac{\partial \mu}{\partial X_i} > \right] P'_{\mu\nu} + \frac{\partial V_{NN}}{\partial X_i} \qquad (26)$$

with $\underline{P}$ and $\underline{P}'$ as defined in Eq. (10).

We must now examine the integrals that appear in Eq. (26). We first examine the derivatives of an atomic orbital with respect to the displacement of the nucleus on which it is centered. Most semi-empirical methods use Slater type orbitals (STO's) as defined in Eq. (27):

$$X(n\ell m) = N_{n\ell m}\, r_a^{n-1}\, \ell^{-\zeta r_b}\, P_\ell^m(\cos\theta)\, \ell^{im\phi}$$

$$N_{n\ell m}^2 = (2\zeta)^{2n+1} \left[\frac{(2\ell+1)(\ell-|m|)!}{(2n)!\ 4\pi(\ell+|m|)!} \right] \qquad (27)$$

Using the cosine law, and the relationship (25)

$$\frac{\partial P^m(\cos\theta)}{\partial\cos\theta} = (\cos^2\theta - 1)^{-1}[\ell\cos\theta\ P_\ell^m(\cos\theta) - (\ell+m)P_{\ell-1}^m(\cos\theta)] \tag{28}$$

and

$$\frac{\partial r_a}{\partial R_A} = \cos\theta \quad , \tag{29}$$

one obtains after quite a bit of algebra

$$\frac{\partial X(n\ell m)}{\partial R_B} = \frac{\zeta}{(2\ell+1)^{\frac{1}{2}}} \Bigg\{ 2(n-1-\ell) \left[\frac{(\ell + 1 + |m|)(\ell \quad 1 - |m|)}{2n(2n-1)(2\ell+3)}\right]^{\frac{1}{2}}$$

$$\times \quad \chi(n-1,\ell+1,m) + 2(n+\ell) \left[\frac{(\ell + |m|)(\ell - |m|)}{2n(2n-1)(2\ell-1)}\right]^{\frac{1}{2}} \chi(n,\ell-1,m)$$

$$- \left[\frac{(\ell - |m| + 1)(\ell + |m| + 1)}{(2\ell + 3)}\right]^{\frac{1}{2}} \chi(n,\ell+1,m)$$

$$- \left[\frac{(\ell + |m|)(\ell - |m|)}{(2\ell - 1)}\right]^{\frac{1}{2}} \chi(n,\ell-1,m) \Bigg\} \tag{30}$$

an expression first given by Garrett and Mills.[26] Most *ab initio* calculations are performed using Cartesian Gaussian functions

$$G(\ell mn) = N_{\ell mn}\ X_a\ Y_a^m\ Z_a^n\ \ell^{-\alpha r_a^2}$$

where

$$N_{\ell mn} = \left[\frac{(8\alpha)^{\ell+m+n}\ (\ell-1)!(m-1)!(n-1)!}{(2\ell-1)!(2m-1)!(2n-1)!}\right]^{\frac{1}{2}} \left(\frac{\alpha}{2\pi}\right)^{\frac{3}{4}} \tag{31}$$

The derivatives of $G(\ell mn)$ with respect to the nuclear coordinate on which G is centered is relatively straightforward, yielding

$$\frac{\partial G(\ell mn)}{\partial X_A} = [(2\ell+1)\alpha]^{\frac{1}{2}} G(\ell+1,m,n) - 2\ell \left[\frac{\alpha}{2\ell-1}\right]^{\frac{1}{2}} G(\ell-1,m,n) \tag{32}$$

with similar expressions for $\partial G/\partial Y_A$ and $\partial G/\partial Z_A$. It is understood that in the normalizer $N_{\ell mn}$ of Eq. (31) that $(-1)!/(-1)! \equiv 1$, and that the second term of Eq. (32) is not used when $\ell=0$. Although it appears that Eq. (32) is simpler than Eq. (30), it must be recalled that there are, indeed, separate evaluations for $\partial G/\partial X_A$, $\partial G/\partial Y_A$ and $\partial G/\partial Z_A$ while $\partial \chi/\partial X_A$ etc., are simply obtained from

$$\frac{\partial \chi}{\partial X_A} = \frac{\partial \chi}{\partial R_A} \frac{\partial R_A}{\partial X_A} = \sin\theta_A \cos\phi_A \frac{\partial \chi}{\partial R_A} ,$$

$$\frac{\partial \chi}{\partial Y_A} = \sin\theta_A \sin\phi_A \frac{\partial \chi}{\partial R} , \tag{33}$$

$$\frac{\partial \chi}{\partial Z_A} = \cos\theta_A \frac{\partial \chi}{\partial R} .$$

The derivatives of all one-center integrals are zero, for it has been assumed that the orbitals on center A follow the displacement of center A. The kinetic energy operator and the electron-electron repulsion operator r_{12}^{-1} are not functions of nuclear coordinates. The derivations of the nuclear-nuclear repulsion energy V_{NN} is given simply by

$$\frac{\partial V_{NN}}{\partial X_A} \equiv \frac{\partial}{\partial X_A} \frac{1}{2}\left[\sum_{B,C}{}' \frac{Z_B Z_C}{R_{bc}}\right] = -Z_A \sum_B \frac{Z_B (X_B - X_A)}{R_{AB}^3} \tag{34}$$

and the derivative of the nuclear-electronic attraction term by

$$\frac{\partial V_{NE}}{\partial X_A} \equiv \frac{\partial}{\partial X_A} - \sum_B \frac{Z_B}{r_{ab}} = \frac{Z_A(X_a - X_A)}{r_{aA}^3} \quad . \qquad (35)$$

The above equations of this section are sufficient to calculate the gradients of all integrals, and thus to evaluate Eq. (11) assuming Eq. (9) is valid.

In practice, semi-empirical all-valence electron methods that are in wide use today involve the evaluation of overlap integrals and certain two-electron two-center integrals of the form $(s_A s_A | s_B s_B)$. The derivatives of the overlap are quickly taken using Eq. (30). The two-center integrals, if integrated over STO's, can also be taken using Eq. (30). These integrals, however, can also be expanded as a function of R_{AB},[27,28] and the derivatives are most easily taken directly on these closed expressions. Often semi-empirical methods utilize formula of the type[29,30,31]

$$\gamma_{AB} \equiv (s_A s_A | s_{BB}) = [a^n + R_{ab}^n]^{-1/n} \qquad (36)$$

where

$$a = \frac{2}{(\gamma_{AA} + \gamma_{BB})}$$

and therefore

$$\frac{\partial \gamma_{AB}}{\partial X_A} = -[a^n + R_{AB}^n]^{-\left(\frac{1}{n}+1\right)} R_{AB}^{n-2}(X_A - X_B). \quad (37)$$

The new MNDO method[32] uses multipolar expansions for the NDDO type integrals. Equation (36) is used with n=2 for integrals of the type $(s_A s_A \; s_B s_B)$. For the general integral

$$(\mu_A \nu_A | \sigma_B \lambda_B) = \sum C_{i0} \left[(R_{AB} + C_{i1} \; D_i^A + C_{i2} \; D_i^B)^2 + (C_{i3} \; D_i^A + C_{i4} \; D_i^B)^2 + a^2 \right]^{-\frac{1}{2}}$$

where C_{ij} are constants depending on the type of two-center integral, while D_i^K depends on atom K and represents the distance from nuclei to "point charge".[33] The derivative of such an analytic function of R_{AB} is again straightforward.[34]

The above equations are complete for most semi-empirical methods. The derivative of three- and four-center integrals required in *ab initio* methods for use in Eq. (26) are applications of Eq. (32) and have been worked out and applied by Schlegel and Wolfe.[35,36] Again, it is easier and more effective to take the explicit derivatives of these integrals after they have been expanded as functions of R.

FORCE CONSTANTS

From the above considerations it is clear that force constants can be obtained from the steps utilized in searching the potential energy surface for extrema. If the rotations and translations have been separated from the search, and Type (c) algorithms have been used, then inversion of $\underline{S}_n$ of Eq. (21) should approximate the $\underline{H}$ matrix. A problem with this procedure is that the geometry optimization may have terminated before $\underline{S}_n$ has accurately converted to $\underline{H}^{-1}$. As far as I know, the accuracy of this procedure has not been checked.

Another procedure is to calculate the second derivatives numerically from the first derivatives at the extreme point. If the first derivatives are analytic, then second derivatives can be obtained from[38,39]

$$H_{ij} = \frac{\partial^2 E}{\partial X_i \partial X_j} = \tfrac{1}{2}\left\{ \frac{f_i(\underline{X}_e + \alpha_j \underline{e}_j) - f_i(\underline{X}_e - \alpha_j \underline{e}_j)}{\alpha_j} + \frac{f_j(\underline{X}_e + \alpha_i \underline{e}_i) - f_j(\underline{X}_e - \alpha_i \underline{e}_i)}{\alpha_i} \right\} . \qquad (38)$$

Both terms that appear in Eq. (38) should be equal, and their difference is a measure of the accuracy of the numerical second differentiation. If the rotations and translations have not been factored then the accuracy of this procedure can be further checked by the number of significant figures that the six zero eigenvalues of $\underline{H}$ are truly zero.

A generalization of Eq. (24) can also be used to obtain the elements of $\underline{H}$;[37] i.e.,

$$\alpha H_j = \mu\delta f_j(0) - \frac{\mu\delta f_j^3(0)}{3!} + \ldots \frac{(-1)^n (n)!^2}{(2n+1)!} \mu\delta f_j^{2n+1}(0) \qquad (39)$$

with

$$\delta f_j(r) = f(\underline{X} + r\alpha\underline{e}_j)$$

$$\mu\delta f_j^{2K+1}(0) = \tfrac{1}{2}(\delta f_j^{2K}(1) - \delta f_j^{2K}(-1))$$

and

$$\delta f_j^{2K}(r) = \delta f_j^{2K-2}(r+1) - 2\delta f_j^{2K-2}(r) + \delta f_j^{2K-2}(r-1)$$

Again, the first term of Eq. (39) is Eq. (38). Note that in the use of Eq. (38) or (39), searches along $\underline{X} \pm r\alpha_j\underline{e}_j$ complete one column of the $\underline{H}$ matrix, $\underline{H}_j$, when all $\underline{f}$ are analytically evaluated (i.e., the first terms of Eq. (38) for all i). In using the simpler form of Eq. (38) only 2m+1 calculations need be performed.

If both first and second derivatives are obtained numerically, then H_{ij} is best obtained by

$$H_{ij} = [E(\underline{X} + \alpha_i\underline{e}_i + \alpha_j\underline{e}_j) + E(\underline{X}) - E(\underline{X} + \alpha_i\underline{e}_i) - E(\underline{X} + \alpha_j\underline{e}_j)] / 2\alpha_i\alpha_j \quad , \qquad (40)$$

an equation easily derived from previous considerations, and utilized by Payne.[20]

One might also consider the evaluation of second derivatives analytically. This is a difficult business, however, not only because of the additional derivatives of the integrals that must be performed, but also because the first order changes to the wave function with respect to nuclear displacement need be considered.

TRANSITION STATES

Following McIver and Komornicki[37] a transition state is defined when $\underline{f}(\underline{X}_e) = \underline{0}$ (a stationary point) and one and only one eigenvalue of $\underline{H}(\underline{X}_e)$ is negative. These two considerations define a simple saddle point. In addition, $\underline{E}(\underline{X}_e)$ should be the highest energy point on a continuous line connecting reactants and products; i.e., $\underline{X}_e$ should represent the saddle point of highest energy. Such a definition tends to associate clearly one side of the "pass" with reactants, the other with products. In addition, $\underline{X}_e$ must represent the lowest energy point which satisfies the above three conditions. Defining, as before

$$\sigma(\underline{X}) = \sum_{i=1}^{m} f_i^2(\underline{X}) = \underline{f}(\underline{X})\ \underline{f}^+(\underline{X}) \tag{41}$$

we seek points in which $\sigma(X_e) = 0$. Since $\sigma(X) \geqslant 0$, least squared minimization procedures are appropriate. Such a procedure, however, will force convergence on any stationary point $\underline{X}_e$, so care must be taken with the guesses on initial geometries. Chemical intuition will be of great use here.

There are many methods of least squared minimization. The general starting point is, again, a Taylor expansion

$$\sigma(\underline{X}_{K+1}) = \sigma(\underline{X}_K) + \underline{q}_{K+1}\ \underline{V}_K^+ + \tfrac{1}{2}\underline{q}_{K+1}\ \underline{W}_K\ q_{K+1}^+ + \cdots \tag{42}$$

$$\underline{q}_{K+1} = \underline{X}_{K+1} - \underline{X}_K \quad ,$$

where V_K^+ is a column vector, the elements of which are

$$V_i^K = \left.\frac{\partial\sigma(\underline{X})}{\partial X_i}\right|_{\underline{X}=\underline{X}_K} , \tag{43}$$

$$W_{ij}^K = \left.\frac{\partial^2\sigma(\underline{X})}{\partial X_i \partial X_j}\right|_{\underline{X}=\underline{X}_K} . \tag{44}$$

At the minimum value of σ, $\sigma(X_e) = 0$, and $\underline{V}(\underline{X}_e) = 0$, suggesting the iterative equation

$$\underline{q}_{K+1} = -\underline{V}_K \underline{W}_K^{-1} . \tag{45}$$

Since is given by Eq. (41)

$$V_K = 2\underline{f}_K\underline{H}_K , \tag{46}$$

and

$$\underline{W}_K = 2(\underline{H}_K\underline{H}_K^+ + \underline{C}_K) ,$$

with

$$C_{ij}^K = \sum_{\alpha=1}^{m} f_m \frac{\partial^2 f_m}{\partial X_i \partial X_j} . \tag{47}$$

σ can be minimized in exactly the same fashion in which E itself was minimized, for example the Murtagh-Sargent procedure already described. A similar algorithm described by Powell[40] has been applied with success by Poppinger.[41] A generalized Newton-Raphson method has been employed by McIver and Komornicki.[37] In their application

$$\underline{W}_K \approx 2H_K H_K^+ . \tag{48}$$

Interestingly, W_K by this construction is guaranteed positive semi-definite, and is a good candidate for the Murtagh Sargent procedure. McIver and Kormornicki, however, suggest taking the inverse of W_K explicitly. If this is to be done, the rotations and translations must be factored from the problem. They also remark that higher order terms in Eq. (39) are desirable in the form of $\underline{H}$; i.e., more accuracy in the formation of $\underline{H}$ lessens the number of cycles required in the calculation of simple saddle points.

CONSTRAINED VARIATION

It is clear from the outset that the fewer degrees of freedom that are varied in the study of the energy surface of a system, the easier the procedure will be to obtain stationary points on that surface. The five or six degrees of freedom representing translation and rotation may always be removed exactly without any real constraints to the optimization procedure. If there exists symmetry in the system, and this symmetry is known to persist throughout the problem of interest, then symmetry-adapted coordinates may be used in the optimization procedure, again simplifying the calculation. Considering formaldehyde, H_2CO, there are 12 coordinates, reduced to 6 by removing translation and rotation. Considering C_{2v} symmetry, only 3 variables remain, the CO and CH bond lengths and the OCH angle.

The above constraints do not affect our ability to obtain exact stationary points on the potential energy surface. To ease the calculation, however, we might also consider constraints on the variables guided by "chemical intuition." In the above example on formaldehyde we might fix the CH bond at a typical value of 1.1Å and vary only the CO bond length and the OCH angle. If we are interested in the biphenyl C-C bond between the phenyl moieties, we might fix all the coordinates except this C-C bond length and the dihedral angle between the two phenyl planes. The savings of effort can be substantial, but it is clear that the accuracy of the results obtained will depend on the accuracy of the starting intuition.

Somewhat more dangerous is the use of such intuition for problems that follow pathways on the surface (valleys). In examining internal rotations as, for example, that of ethane, it is tempting to freeze all bond lengths and angles except that representing the torsional one. This is a reasonably accurate procedure, but if one has started with optimized coordinates for the minimum, the barrier, calculated without all coordinates relaxed, might be too large. In the search for reaction pathways, freezing coordinates will generally lead to an overestimate of barrier energies as our knowledge about minima is far greater than that about transition states.

Worse, by freezing coordinates we prejudice the direction of the path, and so can completely miss alternate pathways, the lowest energy pathways, and perhaps even the most important transition state! A rather interesting example of the sensitivity of some conformations to the relaxation of all coordinates is given by Peterson and Csizmadia[42] in their study on the topology of n-butane. While the anti-conformation was exactly at the point predicted by "intuition," the gauche conformation was sensitive to CH_3 group torsional relaxations.

SOME EXAMPLES

The examples in which geometries have been estimated by quantum chemical calculations are, indeed, numerous. For the purposes of these notes we might show two examples. The first of these is to demonstrate the efficiency of the Murtagh Sargent procedure in obtaining the lowest energy conformation of formaldehyde. This is a small molecule, the results of which are easily summarized in Table 1. The variables in this case are all 12 Cartesian coordinates. In spite of this, there has been no reset of the inverse Hessian matrix ($\underline{H}^{-1}$). At "convergence" then, the average root mean square force $\Gamma\sigma/m$ is 0.0006.

Table 2 summarizes the results of a geometry estimate using Eq. (4) and (5), estimating $\underline{f}$ and $\underline{H}$ via Eqs. (24) and (40) for a saddle point rotamer of formamide, Fig. 2. The rotation angle has been constrained to examine the rotation about the CN bond.[21] Nalewajski found that relaxation of the CN bond length during this rotation lowers the barrier from

TABLE 1. Murtagh-Sargent optimization on H_2CO (INDO).

Cycle	1	2	3	4
CO	1.220	1.2831	1.2545	1.2517
CH	1.090	1.0908	1.1186	1.1197
HH	1.888	1.9082	1.9349	1.9317
E	-25.8471	-25.8486	-25.8541	-25.8542
$\sqrt{\sigma}$	0.1817	0.1894	0.0506	0.0069

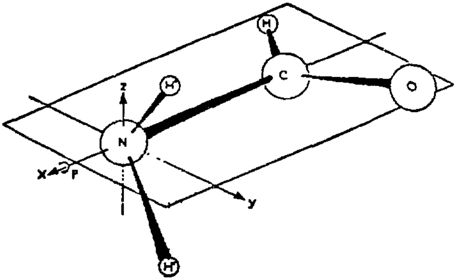


Fig. 2. Rotational conformers of formamide. $\phi \approx 0°$ is the global minimum, $\phi = 90°$ (this figure), and $\phi = 270°$ are two stable rotamers. From Nalewajski, Ref. 21.

TABLE 2. Geometry optimization of a saddle point rotamer of formamide, from Nalewajski.[21]

Variables Optimized	ANGLE (degree) HNH	HNC	HCN	NCO	LENGTHS (Å) NH	CO	CH	CN	ENERGY (A.U.)
Initial Geom.	119.4	120	113.2	123.8	1.010	1.193	1.102	1.376	-168.6405
Optimized:									
4-31G	132.0	114.7	113.5	125.1	1.000	1.196	1.077	1.416	-168.6495
MINDO	101.0	111.6	111.4	122.2	1.147	1.231	1.232	1.439	
Step Size	10.0	10.0	5.0	2.0	0.050	0.020	0.050	0.050	
Gradient+	-0.065	0.082	-0.012	0.021	0.005		0.028	0.064	

+In a.u./Å or a.u./rad.

about 31 kcal/mole to 25 kcal/mole. An interesting aspect of Table 2 is the large step sizes taken to obtain $\underline{f}$ and $\underline{H}$, and the sizes of the gradients calculated at the initial geometry. A nice additional feature would be a recalculation of the forces at the estimated "optimal" geometry to give confidence to the final estimate.

Pulay and coworkers have pioneered work on obtaining force constants and the infrared structure of molecules using the "force field" method.[38,39] Table 3 is the summary of the results obtained by Török, Hegedüs, Kosa and Pulay on one of the fluorinated benzenes they have examined. The results are remarkably good. They have used a simple scaling scheme to correct the CNDO forces that are generally calculated a factor of two greater than observed.[43] Table 4 summarizes a detailed study of the force constants obtained by *ab initio* methods for water. The results for the quadratic force constants, especially for the larger basis sets, is quite satisfactory. The results are generally good for the cubic force constants, and worse for the quartic. In this case, the quartic bending constant $f_{\theta\theta\theta\theta}$ is not satisfactory. These results are obtained through analytic first derivatives of the energy, and numerical estimates of the second derivatives, Eq. (38).

McIver and Kormorniki[37] introduced and demonstrated the utility of least squared minimization techniques for obtaining transition states. An example of the utilization of such a technique is given in Table 5 for the simple rearrangement:

TABLE 3. I.R. frequencies of 1,3-difluorobenzene (cm^{-1}). From Pulay.[38b]

Sym.	Calc. CNDO/2	Exp.	Type via CNDO
A_2	235	251	C–C
	613	599	C–F, C–C
	896	879	C–H
B_1	225	235	C–C
	442	458	C–C
	689	672	C–C, C–F
	804	769	C-H
	902	853	C–H
	1005	978	C–H

TABLE 4. Calculated force constants of water.+

Force Constant	95/41 + 2[a]	STO - 3 G[b]	4-31 G[b]	Exp.[c]
f_{rr}	9.16	10.18	8.71	8.45
f_{rr}'	-0.17	-0.48	-0.13	-0.10
$f_{\theta\theta}$	0.78	1.31	0.79	0.70
$f_{r\theta}$	0.29	0.31	0.32	0.22
f_{rrr}	-60.80	-57.14	-61.40	-59 ± 3
f_{rr}'	0.22	0.76	0.14	0.25 ± 1.5
$f_{rr\theta}$	-0.48	-0.59	-0.00	0.40 ± 0.2
$f_{\theta\theta r}$	-0.41	0.50	-0.43	-0.22 ± 0.1
$f_{\theta\theta\theta}$	-0.86	-1.06	-0.91	-0.88 ± 0.1
f_{rrrr}	437.6	290.2	413.2	384 ± 62
$f_{\theta\theta\theta\theta}$	25.2	-37.9	-30.7	-0.07 ± 0.2

+Units: For stretches mdynes/Å, mdynes/Å^2, etc., for stretch-bond and stretch-bond-bond, mdynes.

a) From Pulay, Ref. 38b.

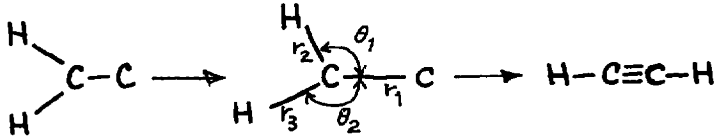
b) From Schlegel and Wolfe, Ref. 35a.

c) From A.R. Hoy, I.M. Mills and G. Strey, Mol.Phys. 24, 1265 (1972).

TABLE 5. Ethenylidine-acetylene rearrangement; from Poppinger.[41]

Variable	Ethenylidine	Transition[a]	Acetylene
r_1(Å)	1.316	1.264 (1.270)	1.178
r_2(Å)	1.086	1.437 (1.270)	2.253
r_3(Å)	1.086	1.084 (1.080)	1.075
θ_1(deg)	121.4	52.8 (60.0)	0.0
θ_2(deg)	121.4	178.2 (160.0)	180.0
Energy (a.u.)	-73.5396	-73.4889 (-73.4883)	-73.6046
DE(kJ/mol)	170.7	303.8	-

a) The numbers in paranthesis give the starting geometry and energy for the transition state, see text.



examined by Poppinger.[41] The starting geometry for the transition state was the symmetric hydrogen bridged structure. Poppinger demonstrated that even relatively poor guesses at the transition state led to the same intermediate as given in the table, but of course the number of calculations required to reach this state is increased.

There are many examples that could be presented. The above, hopefully, are representative of the possible variety.

SOME CONCLUSIONS

Geometry optimization utilizing the axial iteration techniques is reasonably old and straightforward. Some versions of the Gaussian 70 computer program[43] have included this option. Providing the coordinates are not strongly coupled, and a reasonable starting geometry is given, this procedure is successful is reaching minima. The simplex method described begins to have an advantage over the univariant methods when the number of variables increases, insofar as fewer energy calculations are required to reach a stationary point. It should be recalled, however, that univariant searches only require the recalculation of the relatively few integrals involving the coordinate change, while the simplex method requires a recalculation of all integrals. For this reason, the axial procedure was a natural starting point in *ab initio* work, while the simplex method is more effective for the larger molecules of semi-empirical theories.[14]

Methods that assume the potential is quadratic, and build up the gradients are far more effective than either than axial or simplex methods described in the number of SCF calculations required. Of these, the Murtagh Sargent (MS) method[15] outlined seems most effective. Table 6 is an attempt to compare several methods. The results presented for MS with analytic derivatives is an estimate from Poppinger's work[22] recalling that seven calculations are required to obtain the derivatives initially, then six. This estimate was then checked with an INDO program that analytically evaluates the gradients. The number of energy calculations required was, indeed four, when all twelve coordinates were involved (and thus $\underset{\sim}{H}$ is singular). The question then is "can the gradients of the energy be evaluated more rapidly than five or six SCF calculations?" For semi-empirical theories, the gradients are evaluated much more quickly than the SCF step. In *ab initio* calculations the gradients require about the same time as does the SCF.[36,43] At the SCF level then, methods that use analytic gradients are to be preferred over those that do not.

TABLE 6. A comparison of optimization methods: number of energy evaluations.*

Derivatives	No Derivatives			Numerical		Analytic
	Simplex[e]	AIT[a]	DSC[b]	MS[c]	FL[d]	MS[f]
CH_2O	12(3)	28(3)	28(3)	25(3)	20(3)	4(12)
$C_3H_5^+$		199(8)[g]			76(8)	4(20)

* Numbers in parentheses are number of independent variables considered.

a) Ref. 22a, Axial Iteration Technique.

b) Ref. 22a, a variant of the axial iteration technique that allows a change in direction.

c) Murtagh Sargent, Ref. 22a.

d) Fletcher, from Ref. 22a.

e) Nelder and Mead, this work.

f) Murtagh Sargent with analytic derivatives, this work.

g) Reported of lower accuracy, Ref. 22a.

As we have seen, with a reasonably good starting point, two SCF calculations are all that are required if the first and second derivatives are available. The analytic evaluation of the second derivatives, however, is difficult, requiring information on the first order change of the molecular orbital coefficients with respect to geometry. Nevertheless, the elegance of such a procedure is appealing. For semi-empirical methods these derivatives may still be evaluated rapidly. In practice, four or five calculations of the Murtagh Sargent type are required if the analytic first derivatives are available and if we are in the quadratic region of the minimum with our initial guess. Then can the second derivatives be evaluated in the time of four or five SCF-plus-first-derivative calculations? If so, such methods are desirable as the most effective in yielding geometries of molecules at minima and transition states. As an additional bonus, knowledge of the second derivatives yield directly force constants and the infrared spectrum. Initial investigations in this direction seem very encouraging.[44,45,46]

REFERENCES

1. J.A.Pople, in Computational Methods for Large Molecules, F.Herman, A.D.McLean and R.K.Nesbet, editors (Plenum Press, 1973).
2. a) R.C.Bingham, M.J.S.Deum and D.H.Lo, J. Am. Chem. Soc. 97, 1285 (1975);

b) W.A.Lathan, W.J.Hehre and J.A.Pople, J. Am. Chem. Soc. 93, 808 (1971).

3. M.Eliashevich, Compt. Rend. Acad. Sci. URSS 28, 605 (1940); E.B.Wilson,Jr., J. Chem. Phys. 9, 76 (1941).

4. E.B.Wilson, J.C.Decius, P.C.Cross, Molecular Vibrations (McGraw-Hill, 1955).

5. H.Helman, Ein Fuhring in die Quartenchemie (Franz Deuticke, Leipzig, 1937); R.P.Feynmann, Phys. Rev. 41, 721 (1939).

6. A.C.Hurley, Proc. Roy. Soc. A226, 170 (1954); 179 (1954).

7. For example, the floating spherical Gaussian method of Frost yields a wavefunction that does obey the Helman-Feynmann theorem.

8. M.Zerner, unpublished.

9. P.R.Adby and M.A.H.Dempster, Introduction to Optimization Methods (Chapman & Hall, London, 1974); A.Ralston, A First Course in Numerical Analysis (McGraw-Hill, 1965).

10. G.R.Walsh, Methods of Optimization (John Wiley, 1975); Numerical Methods for Unconstrained Optimization, W.Murray, editor (Academic Press, 1972).

11. W.Spendley, G.R.Hext and F.R.Himsworth, Technometrics 4, 441 (1962).

12. a) W.I.Zangwill, Comp. Jour. 7, 293 (1964);
b) M.J.D.Powell, Comp. Jour. 7, 155 (1964).

13. J.A.Nelder and R.Mead, Comp. Jour. 7, 308 (1965).

14. A.Brown, M.J.S.Dewar and W.W.Schoeller, J. Am. Chem. Soc. 92, 5516 (1970).

15. B.A.Murtagh and R.W.H.Sargent, Comp. Jour. 13, 185 (1970).

16. W.C.Davidon, AEC Research & Development report ANL-5990 (1959); Comp. Jour. 10, ,406 (1968).

17. R.Fletcher and M.J.D.Powell, Comp. Jour. 6, 163 (1963)

18. M.Zerner and A.Bacon, unpublished.

19. B.A.Murtagh and R.W.H.Sargent (1968).

20. P.W.Payne, J. Chem. Phys. 65, 1920 (1976).

21. R.F.Nalewajski, J. Am. Chem. Soc. 100, 41 (1978).

22. D.Poppinger, Chem. Phys. Lett. 34, 332 (1975); Chem. Phys. 12, 131 (1976).

23. H.Jeffreys and B.S.Jeffreys, Methods in Mathematical Physics (Cambridge University Press, 1942), p.277.

24. J.A.Pople, J.W.McIver,Jr., and N.S.Ostlund, J. Chem. Phys. 49, 2960 (1968).

25. Handbook of Mathematical Functions, M.Abramowitz and I.A.Stegun, editors (Dover, 1964), p.334.

26. J.Gerratt and I.M.Mills, J. Chem. Phys. 49 (1968).

27. C.C.J.Roothaan, J. Chem. Phys. 19, 1445 (1951).

28. The overlap integrals can be written in terms of A and B functions, and the derivatives of these auxiliary functions are easily taken: R.S.Mulliken, C.A.Rieke, D.Orloff and H.Orloff, J. Chem. Phys. 17, 1248 (1949).

29. K.Ohno, Theoret. Chim. Acta 3, 219 (1964).

30. G.Klopman, J. Am. Chem. Soc. 86, 4550 (1964).

31. N.Mataga and K.Nishimoto, Z. Physik. Chem. (Frankfurt) 13, 140 (1957).

32. M.J.S.Dewar and W.Thiel, J. Am. Chem. Soc. 99, 4899 (1977); 99, 4907 (1977).

33. M.J.S.Dewar and W.Thiel, Theoret. Chim. Acta 46, 89 (1977).

34. M.J.S.Dewar and Y.Yamaguchi, Comp. & Chem. 2, 25 (1978).

35. H.B.Schlegel and S.Wolfe, J. Chem. Phys. 63, 3632 (1975); H.B.Schlegel, Ph.D. Thesis, Dept. of Chemistry, Queens University (1975).

36. H.B.Schlegel, S.Wolfe and F.Bernardi, J. Chem. Phys. 67, 4181 (1977); ibid, 4194 (1977).

37. I.W.McIver,Jr. and A.Komornicki, J. Am. Chem. Soc. 94, 2625 (1972).

38. P.Pulay, Mol. Phys. 17, 197 (1969); 18, 473 (1970); 21, 321 (1971).

39. P.Pulay and F.Török, Mol. Phys. 25, 1153 (1973); W.Meyer and P.Pulay, J. Chem. Phys. 56, 2109 (1972).

40. M.J.D.Powell, Comp. Jour. 7, 303 (1965)

41. D.Poppinger, Chem. Phys. Lett. 35, 550 (1975).

42. M.R.Peterson and I.G.Csizmadia, J. Am. Chem. Soc. 78, 6911 (1978).

43. F.Török, A.Hegedüs, K.Kósa and P.Pulay, J. Molec. Struc. 32, 93 (1976).

44. B.Schlegel, private communication.

45. S.Wolfe, private communication.

46. J.A.Pople, H.Krishnan, H.B.Schlegel, and J.S.Binkley, to be published.

THE CONFIGURATION INTERACTION METHOD

- $H_{op}\Psi = E\Psi$

  (H_{op}: Born-Oppenheimer Hamiltonian)

- Expand electronic wavefunction is an n-particle basis set $\{\Phi_I\}$,

$$\Psi = \sum_I C_I \Phi_I \quad , \quad \langle\Phi_I|\Phi_J\rangle = \delta_{IJ}$$

- Φ_I: linear combination of Slater determinants, $|\phi_1(1)\eta(1)\ldots\phi_n(n)\eta(n)|$

- Expand spatial orbitals in one-particle basis set $\{\chi_p\}$,

$$\phi_i = \sum_p c_p^i \chi_p \quad , \quad (\phi_i|\phi_j) = \delta_{ij}$$

- Variational principle

$$\frac{\partial}{\partial C_I} \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = 0 \quad \text{for all } C_I$$

- $\underline{H}\ \underline{C} = E\ \underline{C}$

  $H_{IJ} = \langle\Phi_I|H_{op}|\Phi_J\rangle$

- If $\{\phi_i\}$ and $\{\Phi_I\}$ are complete, then Ψ and E are exact solution of Schrödinger equation.

- Otherwise $E \geqslant E_{exact}$

- Usually $\Delta E = E - E_{exact}$ is much greater than energy differences of chemical interest.

- Successful CI calculations must rely on cancellation of errors.

- Calculation giving the lowest energy is not necessarily the best!

THREE STEPS OF CI CALCULATION

1. Selection of basis functions $\{\chi_p\}$
2. Construction of orbital basis $\{\phi_i\}$
3. Selection of configurations

There are more variable parameters than in a semi-empirical calculation!

HOW DO WE GET MEANINGFUL RESULTS?

- Convergence of calculated properties with respect to systematic improvements of basis sets.

- Often requires qualitative understanding of the problem at hand.

- Requires a great deal of care to insure one's qualitative understanding is correct.

- Agreement with experiment without convergence is meaningless.

DEFICIENCY OF RHF METHOD

- Incorrect formal behavior for dissociation, curve crossing, united-atom limit, etc., e.g., H_2 at separated atom limit:

$$1\sigma_g^2 \Rightarrow \tfrac{1}{2}\, 1s_a^2 + \tfrac{1}{2}\, 1s_b^2 + \frac{1}{\sqrt{2}}\, 1s_a 1s_b \quad ;$$

He_2 at united atom limit:

$$1\sigma_g^2\, 1\sigma_u^2 \Rightarrow 1s^2 2p^2$$

- Neglect of near degeneracy effects, e.g., C atom $2s^2 2p^2 + 2p^4$; RHF gives poor splitting between ^3P, ^1S, ^1D.

- Neglect of dynamic correlation effects.

- The first two may be remedied by MCSCF method(?)

- The last one is best treated with CI.

THE MCSCF METHOD

- $\Psi = \sum_I C_I \Phi_I \quad , \qquad (\Phi_I | \Phi_J) = \delta_{IJ}$

- $\phi_i = \sum_p c_p^i \chi_p \quad , \qquad (\phi_i | \phi_j) = \delta_{ij}$

- $\dfrac{\partial}{\partial C_I} \dfrac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = 0$ for all I

- $\dfrac{\partial}{\partial c_p^i} \dfrac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = 0$ for all i and p

- i.e., orbitals are also determined variationally.

Uniqueness of MCSCF Orbitals

- For some choice of configurations $\{\Phi_I\}$ the MCSCF method can only determine certain subsets of the occupied orbitals $\{\phi_i\}$ to within an arbitrary rotation.

- This occurs when $\{\Phi_I\}$ is closed with respect to a rotation of $\{\phi_i\}$.

- For a complete CI, MCSCF only serves to partition the 1-particle space into occupied and virtual subspaces.

- The lack of uniqueness may be eliminated by discarding certain configurations from $\{\Phi_I\}$ without loss of generality.

An Example

- $\Psi_{MCSCF} = c_1 1\sigma^2 + c_2 2\sigma^2 + c_3 1\sigma 2\sigma$

  $1\sigma 2\sigma \equiv \frac{1}{\sqrt{2}} (|1\sigma\alpha 2\sigma\beta| - |1\sigma\beta 2\sigma\alpha|)$

- $1\sigma = \cos\theta\ 1\sigma' + \sin\theta\ 2\sigma'$

  $2\sigma = -\sin\theta\ 1\sigma' + \cos\theta\ 2\sigma' \Rightarrow (i\sigma'|j\sigma') = \delta_{ij}$

- $\Psi_{MCSCF} \;=\; c_1' 1\sigma'^2 + c_2' 2\sigma'^2 + c_3' 1\sigma' 2\sigma'$

$$c_1' = c_1\cos^2\theta + c_2\sin^2\theta + \frac{1}{\sqrt{2}}\, c_3\sin 2\theta$$

$$c_2' = c_1\sin^2\theta + c_2\cos^2\theta + \frac{1}{\sqrt{2}}\, c_3\sin 2\theta$$

$$c_3' = \frac{1}{\sqrt{2}}\,(c_1 - c_2)\,\sin 2\theta + c_3\cos 2\theta$$

- Any one of c_i' may be set to zero by appropriate choice of θ.

- Therefore MCSCF calculation with any two of the three configurations would give the same wavefunction. The occupied orbitals would be well defined, but different in each case.

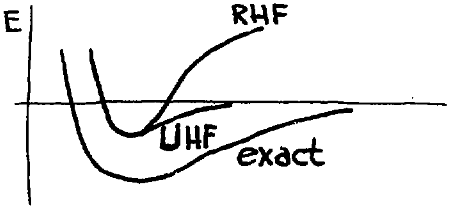
PROPER BOND – DISSOCIATION

- Usually means product of RHF wavefunctions for the fragments at the dissociation limit.

- Examples:

  (1) SPHF (UHF) for LiH $1\sigma\alpha 1\sigma'\beta 2\sigma\alpha 2\sigma'\beta$

  (SPHF = "spin polarized HF")



(2) 2-configuration SCF for LiH $1\sigma^1 2\sigma^2 + 1\sigma^2 2\sigma 3\sigma$

A similar problem to UHF for more complex systems (more on this later).

(3) Complete valence CI MCSCF for LiH $1\sigma^2 2\sigma^2 + 1\sigma^2 3\sigma^2 + 1\sigma^2 2\sigma 3\sigma$; too many configurations.

H_2, THE SIMPLEST CASE

- RHF, $1\sigma_g^2$
- Dissociation limit, $\frac{1}{\sqrt{2}}(1s_a\alpha 1s_b\beta - 1s_b\beta 1s_a\alpha)$
- Form molecular orbitals,

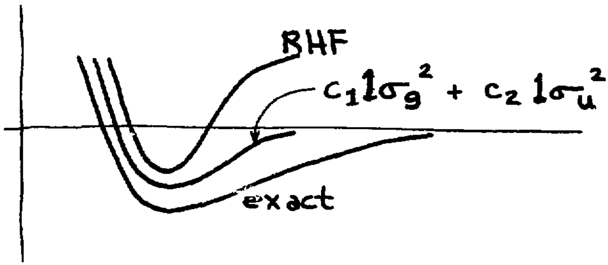
$$1\sigma_g = \frac{1}{\sqrt{2}}(1s_a + 1s_b) \quad , \quad 1\sigma_u = \frac{1}{\sqrt{2}}(1s_a - 1s_b)$$

- Inverse transformation

$$1s_a = \frac{1}{\sqrt{2}}(1\sigma_g + 1\sigma_u) \ , \qquad 1s_b = \frac{1}{\sqrt{2}}(1\sigma_g - 1\sigma_u)$$

- Substitution into separated atom wavefunction gives

$$\frac{1}{\sqrt{2}}(1\sigma_g^2 - 1\sigma_u^2)$$



<u>LiH</u>

- RHF, $1\sigma^2 2\sigma^2$

- Dissociation limit, $\frac{1}{\sqrt{2}}\,(2s_{Li}\alpha 1s_H\beta - 2s_{Li}\beta 1s_H\alpha)$

- Form molecular orbitals,

$$2\sigma = \cos\theta\ 2s_{Li} + \sin\theta\ 1s_H$$

$$3\sigma = -\sin\theta\ 2s_{Li} + \cos\theta\ 1s_H$$

- Inverse transformation,

$$2s_{Li} = \cos\theta 2\sigma - \sin\theta 3\sigma$$

$$1s_H = \sin\theta 2\sigma + \cos\theta 3\sigma$$

- Substitute into dissociation limit wavefunction,

$$\frac{1}{\sqrt{2}}\sin 2\theta\,(2\sigma^2 - 3\sigma^2) + \cos 2\theta\ 2\sigma 3\sigma$$

- Any choice of θ gives correct dissociation limit.

- MCSCF wavefunction independent of θ.

- To dissociate a single bond $\sigma^2 \rightarrow \sigma^{*2}$ is an over-simplification.

- A more useful (but vague) definition: a wavefunction that gives qualitatively correct behavior for the entire dissociation process.

- This would exclude: SCHF and [illegible] product at RHF fragments.

- Complete valence CI still good.

<u>CH $^2\pi$</u>

- RHF, $1\sigma^2 2\sigma^2 3\sigma^2 1\pi$

- Dissociation limit wavefunction in AO,

$$\frac{2}{\sqrt{6}}\ 2p_\sigma\alpha 2p_\pi\alpha 1s_\beta - \frac{1}{\sqrt{6}}\ (2p_\sigma\alpha 2p_\pi\beta + 2p_\sigma\beta 2p_\pi\alpha)1s_\alpha$$

- AO $\rightarrow$ MO,

$$2p_\sigma = \cos\theta 3\sigma + \sin\theta 4\sigma$$

$$1s = -\sin\theta 3\sigma + \cos\theta 4\sigma$$

- Substitute into dissociation limit wavefunction

$$\frac{\sqrt{6}}{4}\ \sin 2\sigma\ (-3\sigma^2 + 4\sigma^2)1\pi + \frac{\sqrt{3}}{2}\ \cos 2\theta\ 3\sigma 4\sigma(^1\Sigma^+)1\pi$$

$$+ \tfrac{1}{2}\ 3\sigma 4\sigma(^3\Sigma)1\pi$$

- θ can be chosen to make one of the first three terms vanish without loss of generality.

- The last term always exists
  - electron recoupling term
  - perfect pairing does not dissociate correctly for open shell molecules

- RHF + product of fragments does not go smoothly to dissociation limit,

$$3\sigma^2 1\pi + 3\sigma 4\sigma 1\pi \; (3\sigma 1\pi\, ^3\Pi)$$
$$\downarrow$$
$$c_1 3\sigma 1\pi(^3\Pi)4\sigma + c_2 3\sigma 1\pi(^1\Pi)4\sigma \quad ,$$

(c_1 and c_2 are fixed).

C_2H_4

- RHF, $\sigma^2\pi^2$, consider the bonds only

- Dissociation limit wavefunction in AO

$$\sigma_a\pi_a\, {}^3\Pi \times \sigma_b\pi_b\, {}^3\Pi$$

$$\frac{1}{\sqrt{3}}\sigma_a\alpha\pi_a\alpha\sigma_b\beta\pi_b\beta + \frac{1}{\sqrt{3}}\sigma_a\beta\pi_a\beta\sigma_b\alpha\pi_b\alpha$$

$$- \frac{1}{2\sqrt{3}}(\sigma_a\alpha\pi_a\beta + \sigma_a\beta\pi_a\alpha)(\sigma_b\alpha\pi_b\beta + \sigma_b\beta\pi_b\alpha)$$

- $1a' = \cos\theta'\sigma_a + \sin\theta'\sigma_b$

  $2a' = -\sin\theta'\sigma_a + \cos\theta'\sigma_b$

  $1a'' = \cos\theta''\pi_a + \sin\theta''\pi_b$

  $2a'' = -\sin\theta''\pi_a + \cos\theta''\pi_b$ (set $\theta' = \theta'' = 45°$)

- Dissociated limit wavefunction

$$-\frac{\sqrt{3}}{2}(1a'^2 - 2a'^2)(1a''^2 - 2a''^2) + \tfrac{1}{2}(1a'2a'\,{}^3A' \times 1a''2a''\,{}^3A')$$

C_2H_4

- The recoupling term $1a'2a'\,{}^3A' \times 1a''2a''\,{}^3A'$ cannot be made to vanish by choice of θ' and θ''.
- Perfect pairing does not dissociate correctly for multiple bonds.
- RHF + product of fragments does not go smoothly to dissociation.

OH $^2\Sigma^+$ Dissociation to Excited States

- RHF, $1\sigma^2 2\sigma^2 3\sigma 1\pi^4$
- Dissociation limit wavefunction in AO,

  $1s^2 2s^2 2p^4 \times 1s \Rightarrow c_1 1s^2 2s^2 2p_\sigma^2 2p_\pi^2 1s + c_2 1s^2 2s^2 2p_\pi^4 1s$

- Dissociation limit wavefunction in MO,

  $c_1 1\sigma^2 2\sigma^2 3\sigma 1\pi^2 + c_2 1\sigma^2 2\sigma^2 3\sigma 4\sigma^2 1\pi^2$

He_2^+: Open-Shell Symmetric Molecules

- RHF, $1\sigma_g 1\sigma_u$
- Dissociation, $\rightarrow \frac{1}{\sqrt{2}}(1s_a^2 1s_b - 1s_a 1s_b^2)$

- This is *not* correct dissociation

- Dissociation limit in AO: $\frac{1}{\sqrt{2}}\,(1s_a^2 1s_b^1 - 1s_a^1 1s_b^2)$

- Form orthonormal set,

$$1s' = c_1 1s + c_2 2s \Rightarrow \langle is|js\rangle = \delta_{ij}$$

$$1\sigma_g = \frac{1}{\sqrt{2}}\,(1s_a + 1s_b) \quad , \qquad 2\sigma_g = \frac{1}{\sqrt{2}}\,(2s_a + 2s_b)$$

$$1\sigma_u = \frac{1}{\sqrt{2}}\,(1s_a - 1s_b) \quad , \qquad 2\sigma_u = \frac{1}{\sqrt{2}}\,(2s_a - 2s_b)$$

- Dissociation limit wavefunction in MO

$$\tfrac{1}{2}\left[c_1 1\sigma_g^2 1\sigma_u + c_2 1\sigma_g^2 2\sigma_u + c_2 1\sigma_u^3 2\sigma_u - \sqrt{2}\;c_2\left(1\sigma_g 1\sigma_u\left({}^1\Sigma_u\right) 2\sigma_g\right)\right]$$

- In general, a very difficult problem; too many configurations

BeH

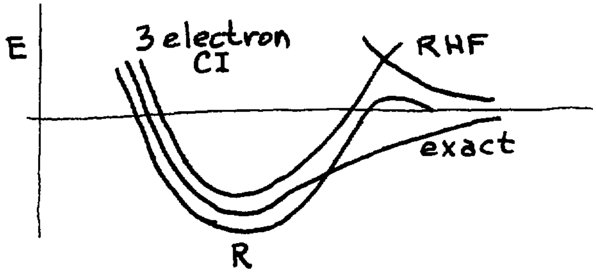
- RHF, $1\sigma^2 2\sigma^2 3\sigma$

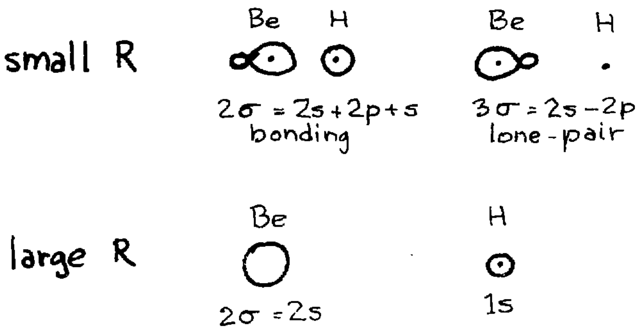
- Apparently correct dissociation,

$$1\sigma \rightarrow 1s_{Be}, \quad 2\sigma \rightarrow 2s_{Be}, \quad 3\sigma \rightarrow 1s_H \ .$$

- Complete 3-electron CI in 2σ and 3σ still gives a maximum at long-range.

- Curve crossing sometimes causes problems.


E
3 electron
CI
RHF
exact
R


small R
Be
H
Be
H
2σ = 2s + 2p + s
bonding
3σ = 2s − 2p
lone-pair
large R
Be
H
2σ = 2s
1s

He_2: Separated and United Atom Limit

- $1\sigma_g^2 1\sigma_u^2 \rightarrow 1s_a^2 1s_b^2$, $R = \infty$

  $\rightarrow 1s^2 2p^2$, $R = 0$

- $1\sigma_g^2 1\sigma_u^2 + 1\sigma_g^2 2\sigma_g^2 \rightarrow$ include ionic terms $R = \infty$

  $\rightarrow 1s^2 2s^2$, $R =$.

- Therefore, short of complete CI, it is difficult (impossible?) to write down a wavefunction with correct behavior at both limits.
- Another example of curve crossing.

PROPER DISSOCIATION

- Correct dissociation limit
- *Insure correct dissociation behavior*, $\sigma^2 \rightarrow \sigma^{*2}$
- Electron recoupling terms
- Look out for complications caused by curve crossing, etc.

COMPLETE VALENCE CI USING MCSCF ORBITALS

- Gives qualitatively correct PES for reactions. (PES = potential energy surface)
- Good for bond lengths, usually too long.
- Good for frequencies, ~100 cm^{-1}.
- Poor for dissociation energies ~1 eV.
- Excitation energies not quantitative.

EMPIRICAL POTENTIALS, SEMIEMPIRICAL POTENTIALS, AND MOLECULAR MECHANICS

Lecture 16/18

by

Norman L. Allinger
Department of Chemistry
University of Georgia
Athens, Georgia

MOLECULAR MECHANICS

Molecular mechanics is a method for studying various physical properties of molecules. (It can sometimes be extended to certain chemical properties, and to smaller and larger systems such as atoms and crystals.) It does not require solution of the Schrödinger equation for the electronic system. In fact, no *explicit* considerations of electrons are required in the usual case. Rather, we consider Van Der Waals interactions between atoms – which involves all the electron-electron, nucleus-nucleus, and nucleus-electron interactions in a simple empirical way. Molecular mechanics also includes the interactions between instantaneous dipoles (electron correlation).

What can we do with it in practice? Determine

1) molecular structure
2) energies: isomerization, conformational, heat of formation
3) vibrational spectra, thermodynamic functions.

ADVANTAGES

Fast and Accurate. Compared to *ab initio* calculations on a molecule containing 3 to 4 first row atoms at, say, the 4-31G level of accuracy, the computer time required is less by perhaps 10^3. If the molecule is larger, the advantage increases rapidly (time approximately N_a^2 vs. N_o^4 where N_a is number of atoms, N_o is number of orbitals). For atoms heavier than first row, computational time is the same as for light atoms by molecular mechanics, but N_o^4 still holds for *ab initio* methods. Molecular mechanics includes the effects of correlation energy in an approximate way. Of course, the 4-31G level gives results that are inadequate for many purposes, in which case, a larger basis set (and much more computer time) would be required for the *ab initio* work.

DISADVANTAGES

Empirical parameters must be known. For molecules such as hydrocarbons, one has 98% of the data one would like, and 95% of it is correct. As refinements continue, there is little that cannot be done accurately to give structures and energies of hydrocarbons. Vibrational spectra and thermodynamic functions are treated less well, but generally better than by *ab initio* methods.

On the other hand, for functionalized molecules (most of them), one has perhaps 85% of the structural information needed, and 50% of the energy information. Overall, the reliability is much less good, although for many restricted classes of molecules, it approaches the hydrocarbon reliability.

Why does a quantum chemist want to use molecular mechanics?

Probably to calculate *structures*. In most cases of interest, one can obtain structures of "experimental quality" by molecular mechanics in a day or two (compared to, say, a month by crystallography). Best of all, you do not need a sample of the compound!

Except for very simple molecules, one cannot optimize all internal degrees of freedom by *ab initio* calculations, so if a structure is needed as a starting point, and a reliable experimental structure is unavailable, the molecular mechanics procedure offers a quick easy way to get a structure, *if* the necessary parameters are available.

Energies. Again, these can be well calculated (competitive with experiment) for hydrocarbons, pretty well for several classes of functionalized compounds, but lots of classes of compounds have not been studied yet, and the necessary parameters do not exist.

Vibrational spectra. Few force fields have so far considered spectra. The results here are much more sensitive to parameters, and the results are usually on the order of $\pm 20\ cm^{-1}$ with hydrocarbons, and are not expected to improve for functionalized molecules (there are few studies to date).

There are a great many force fields in the literature. They are constructed in ways that are generally the same but differ some in detail. Depending on what kinds of things the authors were interested in, different experimental facts were incorporated in the different programs, and consequently, the different programs give different predictions, although the differences are small, on the whole. The following table contains my own assessment of the capabilities of many of the currently available force fields. (My assessment is based only on the information I have available, and may not be completely accurate; but it's the best I can do.) It is offered only as a guide to the uninformed user who may wish to choose between various possible force fields for different kinds of problems.

In the program to be discussed and used here (MM1), the stretching energy for each bond in the molecule is given by $E_s = \frac{k}{2}(\ell-\ell_o)^2(1 + C_s(\ell-\ell_o))$, where C_s has a fixed numerical value for all bonds.

The value of k differs for different kinds of bonds (C-H, C-C, C-O, etc.), as does the value of ℓ_o. These values were all established by studies on simple molecules where sufficient experimental data exist to permit their evaluation. Lists of these numerical values can be obtained from the MMI program, and updates are given in the manual (available from QCPE).

For bending, a similar function is used:

$$E_b = k(\theta-\theta_o)^2(1 + C_f(\theta-\theta_o))$$

and the constants have similar meaning with respect to bending and were evaluated as were the stretching constants above.

The Van Der Waals interactions proved to be quite difficult to quantify. Theory and experiment agree that at longer distances, two rare gas atoms have an attraction between

TABLE 1. Limitations of some popular force fields.

		Saturated Hydrocarbons					
Author	Force field	Structure	Energy H_f^0	Vibra. spectra	Thermodyn. functions	Conjunctive hydrocarbons	Functionalized molecules
R.H.Boyd	-	B	(B)	B	B	(C)	-
L.S.Bartell	MUB-1	B^+	-	-	-	-	-
L.S.Bartell	MUB-2*	A	(B)	-	-	-	-
S.Lifson	Ermer-Lifson	B^+	(B)	A	-	-	-
Karplus-Warshel	-	(B)	(B)	A	-	(B)	-
Schleyer	-	B^+	B	-	-	-	-
Allinger	1971	B^-	B^-	-	-	-	C
Allinger	MM1(1973)	B	B	-	C	C^+	C^+
Allinger	MM2(1977)	A	A	-	-	B	B
White	-	A	B	-	-	-	-

Letters A, B, C are relative grades. A is current state of the art, probably not perfect; B is a average; C is semi-quantitative only. Parentheses indicate insufficient data to evaluate with certainty, but a best guess is given.

*With improved torsional terms added later to original MUB-2.

them which lowers their energy by r^{-6}. This energy (due to electron correlation), and called London or dispersion energy is usually put in the form

$$E = -a\left(\frac{r_o}{r}\right)^6$$

The minus sign means the energy goes down as the atoms approach, and the distance is expressed in units of the sum of the Van Der Waals energy of the atoms involved. The parameter a is empirically adjusted to give the correct magnitude of E at any one distance r. How all this applies to atoms in molecules is not obvious. What has been done is to adjust r_o and a empirically to fit known data on molecules. The intramolecular interactions are summed over all atoms which are not bound to each other or to a common atom.

The part of the Van Der Waals function that led to trouble in practice was the repulsive part. As two rare gas atoms approach, the energy first goes down from dispersion, then as they approach still closer, it abruptly goes up very steeply. A Lennard-Jones potential:

$$E = -a\left(\frac{r}{r_o}\right)^6 + b\left(\frac{r}{r_o}\right)^{12}$$

is commonly used to represent the total behavior.

A Buckingham potential

$$E = -a\left(\frac{r}{r_o}\right)^6 + b\,\exp\left(-\frac{cr}{r_o}\right)$$

is indistinguishable from the Lennard-Jones function in the region of interest and is used in MM1.

It is now recognized that the exponent 12 in the Lennard Jones potential (and the corresponding constant d in the Buckingham) is too large, and a value of 9 or 10 is better. The value 12 gives "harder" atoms, and MM2 is better than MM1 in this respect.

TORSION

In ethane the observed torsional barrier is not obtained using only the above three functions. One must add a term of the type

$$E_w = \frac{V_3}{2} (1 - \cos 3\omega) .$$

For unsaturated molecules such as ethylene, a term of the kind

$$E_w = \frac{V_2}{2} (1 + \cos 2\omega)$$

is needed.

These prototype molecules suggested that simple torsional terms as shown would be adequate for saturated and unsaturated molecules, respectively, and MM1 works this way. We now know that a three-term Fourier expansion is needed for good results:

$$E = \frac{V_1}{2} (1 - \cos\omega) + \frac{V_2}{2} (1 + \cos 2\omega) + \frac{V_3}{2} (1 - \cos 3\omega)$$

(The signs of the constants, the signs of the three terms, and the signs in front of $\cos n\omega$ can be written in many different combinations, and there are no generally accepted conventions.)

Such a function is included in MM2 (and was an addend to MUB-2). With MM1, only the single V_3 term was used. This would have resulted in energies which were too low for gauche-butane type interactions, so a hard hydrogen was used to compensate. The resulting force field is good, but with the 3-term function, one can do better. If one uses only the 3-fold term, some error is present which cannot be corrected completely by any method yet found. Schleyer (also White) reduced the *Van Der Waals* interaction between carbon and hydrogen to a very low value, which permits a soft hydrogen and a good gauche-butane energy. Overall, however, the results were not any better than with MM1. Others (Ermer, Bartell, MUB-1 and the original MUB-2) have used a soft hydrogen and accepted the gauche-butane error, so the structures are generally better but the energies poorer than with MM1.

FURTHER IMPROVEMENTS

As bond angles are compressed, the two bonds including the angle are generally lengthened, and the reverse is also true. This led to the early development of the Urey-Bradley force field, in which two atoms bound to a common atom have an optimum distance, and a Hooke's law (usually) relationship also applies at that distance for each pair of atoms bound to a common atom. An alternative is to add to the valence force field, so-called cross terms, of which the ***stretch-bend*** type is the most important in terms of the effect on geometry.

$$E_{\text{stretch-bend}} = \frac{k_{s\theta}}{2} (\ell - \ell_o)(\theta - \theta_o)$$

Other types of cross terms (stretch-stretch, torsion-bend, etc.) are often important for spectral calculations, but not ordinarily for geometry and energy.

Snyder and Schachtschneider showed that using an equivalent number of parameters, a Urey-Bradley force field and a valence force field with cross terms give similar results.

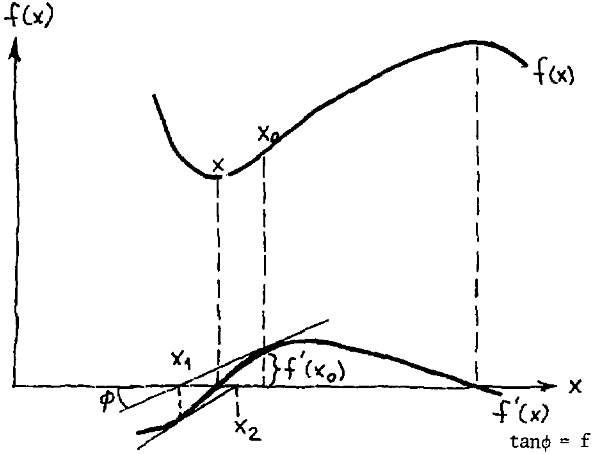
GEOMETRY OPTIMIZATION

The most important feature of force-field calculations is that they are able to take a rough, approximate structure and optimize it to an accurate "experimental" structure. If the functions given above are considered in relation to a molecule, they define a multi-dimensional potential surface, where energy can be imagined as the vertical coordinate. To find the structure one needs to find the location of the energy minimum on the surface. That point gives the (ground state) structure of the molecule and also tells us something about its energy. If a molecule has several conformations, there will be several minima, separated by saddle points. All geometry optimization routines operate by starting from the initial geometry and minimizing the energy (or locating places on the surface where $\partial E/\partial X_i = 0$ for all coordinates (internal or Cartesian) X_i. So if one starts with an approximate structure for, say, gauche-butane, one will obtain an optimized structure for gauche-butane, not for the more stable anti-conformation. Finding the *global* energy minimum is not usually a problem for small or simple molecules, but it can be for large molecules.

ENERGY MINIMIZATION

There are two levels of sophistication that can be used, and a host of variants. Basically, one can use only first derivatives of the energy with respect to the coordinates (Steepest Descent Methods) or one can use both first and second derivatives. These methods are all approximations so one begins with a starting geometry and improves it by successive iterations. The first derivative methods are simple to implement, free from hang-ups and very fast per iteration, relative to the second derivative methods. They are most useful in the early stages of an optimization. As one approaches the energy minimum, the improvement per iteration becomes quite small, and a large number of iterations is required.

Second derivative methods are usually variants of the Newton-Raphson scheme. Here one solves a set of simultaneous equations (3N-6 equation for N atoms), where the coordinates of the atoms are the unknowns. This is usually done by diagonalizing a matrix. This method is most efficient when near the energy minimum. In that case, while each iteration



$$\tan\phi = f''(x_o) = \frac{f'(x)}{\Delta x}$$

$$\Delta x = \frac{f'(x_o)}{f''(x_o)}$$

Fig. 1. Minimization of a function f(x) by the Newton-Raphson procedure. Iteration starting at x_o improves the solution of $f'(x) = 0$ successively to x_1 and x_2, approaching the true solution x.

is time-consuming, the improvement per iteration is very great and few iterations are required.

Finally, one can use a block-diagonal variant of the Newton-Raphson method. Each of these methods has advantages and disadvantages, which depend to some extent on the problem at hand. Most programs use one method or another, and it is only necessary that the user knows what the difficulties are with that particular method, and how to get around them.

The MM1 program uses the block diagonal scheme, which corresponds to optimizing the atomic positions sequentially. The only place where one may have problems is in a molecule where all of the atoms must move cooperatively (as in the pseudorotation of the boat form of cyclohexane for example).

A few programs (mostly older) used numerical calculation of derivatives. This is easy to program, but such programs run very slowly. Faster (by at least a factor of 10) are the analytical methods, but the programming problems are much greater then (MM1 uses analytic calculation of all derivatives).

HEATS OF FORMATION

There are two ways to approach these, depending on what has been done prior to this point. In principle, if one knows the vibrational levels of the molecule, there is some energy (which can be taken as zero) that corresponds to the minimum energy geometry for the rigid vibrationless model. To get the heat of formation one would need to add the zero point energy, and the statistical mechanical energy which comes from populating the higher vibrational levels at room temperature. This has been done in a cursory way by Boyd.

The alternative method, which has been much more highly developed, and works well in practice, has been used by Schleyer, by White, and by us. In this case the geometries of the molecules correspond to experimental geometries at room temperature, and these are defined and parametrized that way. One can similarly define and parametrize the heat of formation calculations, without reference to vibrational levels.

In the first method, one would have to ascertain the value for the zero of energy as mentioned above. This would result from the bonding energy of the molecule. In the second method, the definition proceeds similarly, except in this case since empirical parameters are used for bond energies anyway, one can simply define it to apply at 25°.

The literature abounds with bond-energy schemes, which when used according to the proper recipe, ordinarily do give quite good heats of formation. They fail in special cases where strain in the molecule is not properly taken into account. Clearly, utilizing a molecular mechanics calculation which explicitly gives one strain, and superimposing this upon a bond energy scheme, one would expect to get quite good heats of formation. For a large sample of 42 saturated hydrocarbons containing most of the kinds of crowded and strained molecules for which experimental heats of formation are known, where the average reported experimental error was 0.40 kcal/mole, the MM1 force field gave 0.60 kcal/mol as the standard deviation between the calculated and experimental values. The MM2 force field reduced this to 0.42 kcal/mole. Thus the results for MM1 were good and those for MM2 are excellent. Schleyer's force field and also the one by White also give good results, comparable with those from MM1.

Alkenes can be treated with an accuracy that appears to approach that for alkanes but fewer data are available and some of it is not as accurate, so the results are probably not quite as good on the whole. For functionalized molecules, the quantity of data falls, as does its accuracy, so the overall reliability is somewhat less than for hydrocarbons. Polyfunctional molecules have been studied only to a slight extent and here the reliability of the calculations is probably a great deal less.

Returning again to the statistical mechanical viewpoint, in principle one should calculate separately the rotational and translational contributions to the heat of formation and also add a PV correction, since heats of formation are at constant pressure. These numbers can be evaluated classically, and have a total combined value of 2.4 kcal/mole.

There are two additional quantities that one needs to add to the calculated heat of formation for best results. One is a conformational population term. If we have a single conformation this term is zero, but if the molecule consists

TABLE 2

Alkane Heat of Formation Data[a]

Compound	Wt	H_f^o Calc	H_f^o Exp[b]	Difference (Calc-Exp)	Reported probable errors
Methane	1	−17.82	−17.89	0.07	0.08
Ethane	2	−20.05	−20·24	0·19	0·12
Propane	9	−25·28	−24·82	−0·46	0·14
Butane	9	−30·26	−30·15	−0·11	0·18
Isobutane	9	−32·19	−32·15	−0·04	0·16
Pentane	7	−35·20	−35·00	−0·20	0·16
Isopentane	8	−36·62	−36·92	0·30	0·20
Neopentane	7	−41·06	−40·27	−0·79	0·25
Hexane	6	−40·14	−39·96	−0·18	0·19
Heptane	4	−45·09	−44·89	−0·20	0·19
n-Octane	3	−50·05	−49·82	−0·23	0·20
Hexamethylethane	5	−53·19	−53·95	0·75	0·29
2,3-Dimethylbutane	7	−42·16	−42·49	0·33	0·24
2,2,3-Trimethylbutane	6	−48·81	−48·95	0·14	0·27
Cyclobutane	2	6·16	6·38	−0·22	0·10
Cyclopentane	9	−18·02	−18·30	0·28	0·18
Cyclohexane	8	−30·08	−29·50	−0·58	0·17
Methylcyclohexane	7	−37·02	−36·99	−0·03	0·25
3,3-Diethylpentane	4	−55·41	−55·77	0·36	0·40
1,1-Dimethylcyclohexane	6	−44·02	−43·26	−0·76	0·46
cis-Dimethylcyclohexane	3	−41·73	−41·13	−0·60	0·27
trans-Dimethylcyclohexane	3	−43·06	−42·99	−0·07	0·27
Cycloheptane	7	−28·02	−28·22	0·20	0·26
Cyclooctane	4	−28·96	−29·73	0·77	0·33
Cyclodecane	3	−35·35	−36·29	0·94	1·00
trans-Decalin	3	−43·61	−43·54	−0·07	0·55
cis-Decalin	3	−41·11	−40·45	−0·66	0·55
cis-Hydrindane	4	−29·97	−30·41	0·44	0·47
trans-Hydrindane	4	−30·92	−31·45	0·53	0·50
Norbornane	7	−13·29	−12·40	−0·89	0·40
Cubane	1	149·18	148·70	0·48	1·00
Adamantane	3	−33·34	−32·96	−0·38	0·19
Congressane	1	−37·26	−36·64	−0·62	0·60
Bicyclo(2.2.2)octane	6	−23·81	−23·75	−0·06	0·30
cis-Bicyclo(3.3.0)octane	1	−21·43	−22·30	0·87	0·50
trans-Bicyclo(3.3.0)octane	1	−15·01	−15·90	0·89	0·60
trans-syn-trans-Perhydroanthracene	1	−57·22	−58·32	1·10	1·27
trans-anti-trans-Perhydroanthracene	1	−51·13	−52·93	1·80	1·47
Standard Deviation: 0·60					±0·51
Correlation Coefficient: 0·999					

[a] A few new experimental values became available to us after the data in this table were assembled. The newer values have been included in the table together with the current difference between calculated and experimental values. However, the least squares fitting has not been repeated. If it were to be repeated, very small adjustments in the parameters would be expected, although no significant changes would result.

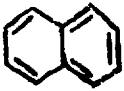
[b] The experimental values are generally taken from Cox and Pilcher, 1970 or API Tables, Project 44, Bureau of Standards, Washington, D.C.

of a mixture of conformations we need to add an increment, which allows for the fact that the higher energy conformations are present in a Boltzmann distribution. Finally, some molecules have torsional vibrations which have small barriers, such as in the case of the torsion about the C-C bonds in alkanes. In other cases the barriers are much higher (as in cyclic compounds). To deal with these simultaneously, it has been found expendient to add a constant, empirically determined, for each bond of the low torsional frequency type. This term is referred to in the MM1 program as a Torsional term.

MOLECULES CONTAINING DELOCALIZED ELECTRON SYSTEMS

Everything up to this point concerns molecules which can be described with a single Kekulé structure. Atoms are either bound together or they are not, and there is no uncertainty on that point. With delocalized electronic systems, however, things are not so simple.

We can perhaps begin by considering two separate cases, butadiene and benzene. In butadiene there are two short bonds and one longer bond. Ordinary polyenes can be treated with parameters which can be picked to fit butadiene. The energies of linear polyenes increase in a linear manner with the number of double bonds and so such systems can be treated in a classical way. However, use of the same numbers will give unsatisfactory results for benzene. It is known that in benzene the bond lengths are all equal, and the energy is a good deal less (resonant) than would be suggested from polyene studies. Again, one can pick different parameters to fit the bonds in benzene, and hence benzene can be dealt with adequately also. Molecules which are either polyenes or simple benzene derivatives can be well treated by the appropriate set of parameters. However, consider what happens if we want to examine a molecule such as naphthalene:



If we use the benzene parameters, we will obtain essentially all equal bond lengths. *If we use the butadiene parameters,* we will obtain bonds which are strongly alternating long and short. Experimentally it is observed that an intermediate situation exists. Howe are we to reproduce that?

Extensive studies on pi-electronic systems over the last 40 years or so have indicated that molecules such as naphthalene need to have a pi calculation carried out quantum-mechanically and superimposed on a sigma calculation if the experimental facts are to be correctly reproduced. One approach is to do exactly this, and Warshel and Karplus have used this direct approach. The sigma system energy is calculated as usual, the pi system energy by a self-consistent field calculation, and the two are summed. The structure is found by minimizing this total energy. An alternative method has been used by us (program MMP1). Here a self-consistent field calculation (actually VESCF) is carried out on the trial pi system and the bond orders are found in the usual way. It is assumed that linear relationships exist between bond order and stretching force constant on the one hand, and between bond order and natural bond length on the other. The linear relationships are established by the examination of simple compounds. These relationships then being known from the trial pi system SCF calculation, one obtains the bond orders, then the force constants and the natural lengths. These in turn are put into the molecular mechanics calculation, which then proceeds in the usual way. If the geometry changes very much, the pi system calculation is repeated, and so is the entire process described above. The system is then brought to self-consistency with respect to the pi and sigma parts, and the energy minimum found as usual.

This scheme is found to work very well in practice and gives good geometries for molecules such as butadiene, benzene and naphthalene. Looking at several hundred bonds for which experimental data are available for comparison, perhaps 90% of them are calculated to within 2 esd of the crystallographic values. Studies on analogous compounds, or in a few cases later, more refined work, suggest to me that almost all of these discrepancies involve experimental rather than calculational errors. I only know of one molecule (18-annulene) where there seems to be a real and serious conflict between what is calculated and what is found crystallographically.

Heats of formation can be calculated utilizing the MM1 procedure, but the results are not completely satisfying. Hence this has never been programmed and the QCPE program should not be used to obtain heats of formation of conjugated systems. Such calculations lack the proper parameters and, while the program will run and give numbers, the numbers are not meaningful.

ELECTROSTATICS IN MOLECULAR MECHANICS

Some force fields have included charge distributions even in saturated hydrocarbons (Lifson), but our own experiences have indicated that this is not necessary insofar as the calculation of energies and structures goes. Similarly, with monofunctional compounds such as ketone, there is no need to explicitly include the electrostatics part of the calculation. However, if one has a compound which contains two or more dipoles, say, for example, 1,2-dichloroethane, then the electrostatic interaction between the dipoles plays an important part in determining the energy of the molecule, the preferred conformation, and it exerts some effect on the structure. How should these electrostatic effects be allowed for?

The most simple approach would seem to be to treat the system as a collection of point charges (ordinarily placed at atomic centers), or to treat it as a collection of point dipoles (ordinarily placed in bonds). As long as the distance between the dipoles is reasonably large relative to the actual charge distribution within the dipole, this procedure is adequate.

The point charge and point dipole approximations ordinarily give very nearly the same results. In a few cases the results differ, but insufficient study has been put into the problem to decide which approximation, if either, is better on the average.

If one assigns bond moments to different kinds of bonds (which is the MM1 approach), or the equivalent in terms of point charges, then one has a first approximation for calculation of the dipole moment of a molecule and for the deformations which occur, and the energy changes which result from interaction of these dipoles.

This is as far as our MM1 program goes. If one wants to go further, one can ascertain atomic charges by quantum mechanical methods, or alternatively, there is a classical scheme, due originally to Smith and Eyring, which we have generalized. This scheme allows for the principle moments of bonds at the outset and then permits each of these to induce in all of the other bonds of the molecule-induced dipoles, and the total final charge distribution is found. This scheme gives us better dipole moments and energies than the simple scheme above. If one is dealing with molecules in solution, one needs to consider the effect of solvation on this charge assembly. In this case the total charge assembly can be approximated by a dipole plus a quadrupole, which can then be solvated according to a scheme originally due to Onsager.

The original approach described above, which was due to

James Jeans, is of marginal accuracy for molecular mechanics purposes. The modified Smith-Eyring method, plus the solvation treatment (due to R.J.Abraham) is better and is perhaps good enough for molecular mechanics purposes.

GENERAL REFERENCES AND REFERENCES FOR APPLICATION OF MOLECULAR MECHANICS CALCULATIONS TO STUDIES OF VARIOUS COMPOUND CLASSES

For reviews, see:

a. N.L.Allinger, Adv. Phys. Org. Chem. 13, 1 (1976). [A discussion of MM1).

b. O.Ermer, Structure and Bonding, Vol. 27, p.161 (Springer-Verlag, New York, 1976).

c. C.L.Altona and D.H.Faber, Topics in Current Chemistry 45, 1 (1974).

d. E.M.Engler, J.D.Andose and P.von R.Schleyer, J. Am. Chem. Soc. 95, 8005 (1973).

3. J.E.Williams, P.J.Stang and P.von R.Schleyer, Ann. Rev. Phys. Chem. 19, 531 (1968).

For more recent work, see

L.S.Bartell, J. Am. Chem. Soc. 99, 3279 (1977) [on MUB-2]

N.L.Allinger, J. Am. Chem. Soc. 99, 8127 (1977). [on MM2]

D.N.J.White and M.J.Bovill, J. Chem. Soc. Perk. Trans. 11, 1610 (1977). [on alkanes and alkenes]

N.L.Allinger and D.Y.Chung, J. Am. Chem. Soc. 98, 6798 (1976) [on alcohols and ethers]

N.L.Allinger and M.J.Hickey, *ibid* 97, 5167 (1975) [alkane thiols and thiaalkanes]

N.L.Allinger and J.Kao, Tetrahedron 32, 529 (1976) [sulfoxides]

a. N.L.Allinger, M.J.Hickey and J.Kao, J. Am. Chem. Soc. 98, 2741 (1976);

b. N.L.Allinger, J.Kao, H.Chang and D.B.Boyd, Tetrahedron 32, 2867 (1976). [disulfides]

A.Y.Meyer and N.L.Allinger, *ibid* 31, 1971 (1975). [halides]

N.L.Allinger and A.Y.Meyer, Tetrahedron 31, 1807 (1975) [alkynes]

N.L.Allinger and H.von Voithenberg, Tetrahedron 34, 627 (1978). [phosphines]

N.L.Allinger and H.Chang, *ibid* 33, 1561 (1977) [carboxylic acids]

N.L.Allinger and M.E.Wuesthoff, Tetrahedron 33, 3 (1977) [electrstatics, dipoles, and induced dipoles]

L.Dosen-Micovic and N.L.Allinger, Tetrahedron 34, 3385 (1978) [preceding electrostatic treatment extended to solutions by solvation calculations]

APPENDIX A

(Taken from R.H.Boyd, J. Chem. Phys. 49, 2574 (1968).)

A METHOD FOR THE CALCULATION OF THE MOLECULAR CONFORMATION OF MINIMUM POTENTIAL ENERGY FROM EMPIRICAL VALENCE FORCE POTENTIAL FUNCTIONS

It is assumed that the potential energy of a molecule made up of N atoms ...ijkℓ..., is known as a function of the atomic positions in terms of valence coordinates, r_{ij}, θ_{ijk}, $\phi_{ijk\ell}$, and δ_{ijkm}, where

r_{ij} = the magnitude of the vector joining ij (i.e., bond length or non-bonded interaction)

θ_{ijk} = angle between the vector joining ji and the vector joining jk (i.e., bond angle)

$\phi_{ijk\ell}$ = angle between the planes of ijk and jkℓ (i.e., bond rotation angle)

δ_{ijkm} = angle between the vector joining jm and the plane ijk (i.e., deformation angle of bond attached to an aromatic ring).

The potential energy is then written as

$$U = \sum_{(ij)} U_{ij}(r_{ij}) + \sum_{(ijk)} U_{ijk}(\theta_{ijk}) + \sum_{(ijk\ell)} U_{ijk\ell}(\cos\phi_{ijk\ell}) + \sum_{(ijkm)} U_{ijkm}(\delta_{ijkm}) \qquad \text{(A1)}$$

where the sums are over each set of interactions considered to be present. For example, if atoms ijk are considered to have an angle (θ_{ijk}) interaction, this is included once in the summation.

The potential energy is then expanded in a power series through quadratic terms about a set of trial coordinates, r^o_{ij}, θ^o_{ijk}, ϕ^o_{ijk}, δ^o_{ijkm} as

$$U = U^o\left(..r^o_{ij}, ..\theta^o_{ijk}, ...\phi^o_{ijk\ell}, ...\delta^o_{ijkm}\right) + \sum_{(ij)} \left.\frac{\partial U_{ij}}{\partial r_{ij}}\right|_{r^o_{ij}} \Delta r_{ij}$$

$$+ \tfrac{1}{2}\sum_{(ij)} \left.\frac{\partial U_{ij}}{\partial r^2_{ij}}\right|_{r^o_{ij}} \Delta r^2_{ij} + \sum_{(ijk)} \left.\frac{\partial U_{ijk}}{\partial \theta_{ijk}}\right|_{\theta^o_{ijkm}} \Delta\theta^2_{ijk}$$

$$+ \tfrac{1}{2}\sum_{(ijk\ell)} \left.\frac{\partial^2 U_{ijk}}{\partial \theta_{ijk}}\right|_{\theta^o_{ijkm}} \Delta\theta^2_{ijk} + \sum_{(ijk\ell)} \frac{\partial U_{ijk}}{\delta \cos\phi_{ijk}} {}_{\cos\phi^o_{ijk\ell}} \Delta\cos\phi_{ijk\ell}$$

$$+ \sum_{(ijk\ell)} \left.\frac{\partial^2 U_{ijk\ell}}{\partial \cos\phi^2_{ijk\ell}}\right|_{\cos\phi^o_{ijk}} \Delta\cos\phi^2_{ijk\ell} + \sum_{(ijkm)} \left.\frac{\partial U_{ijkm}}{\partial \delta_{ijkm}}\right|_{\delta^o_{ijkm}} \Delta\delta_{ijkm}$$

$$+ \tfrac{1}{2}\sum_{(ijkm)} \left.\frac{\partial^2 U}{\partial \delta_{ijkm}}\right|_{\delta^o_{ijkm}} \Delta\delta^2_{ijkm} \quad , \qquad \text{(A2)}$$

where cross terms have been neglected, and

$$\Delta r_{ij} = r_{ij} - r^o_{ij}$$

$$\Delta\theta_{ijk} = \theta_{ijk} - \theta^o_{ijk} \text{ , etc.}$$

Minimizing Eq. (A2) by differentiation with respect to Δr_{ij},

$\Delta\theta_{ijk}$, $\Delta\phi_{ijk\ell}$, and $\Delta\delta_{ijkm}$ is not practical since these coordinates are not independent and the equations of constraint are not easily formulated. However, Δr_{ij}, $\Delta\theta_{ijk}$, $\Delta\phi_{ijk\ell}$, and $\Delta\delta_{ijkm}$ can be transformed to Cartesian coordinates

$$\chi_i^\alpha \qquad (\alpha = 1,2,3; \quad i = 1,\ldots,N)$$

where α refers to the three space coordinates of the i^{th} atom.

The Cartesian coordinates are independent and the transformed potential energy can be minimized by differentiation with respect to χ_i^α. The transformation is made by regarding r_{ij}, $\Delta\theta_{ijk}$, $\Delta\phi_{ijk\ell}$, and $\Delta\delta_{ijkm}$ as small quantities

$$\Delta r_{ij} = \sum_{\alpha=1}^{3} I_r^\alpha \Delta\chi_i^\alpha + \sum_{\alpha=1}^{3} J_r^\alpha \Delta\chi_j^\alpha + \tfrac{1}{2}\sum_{\alpha,\beta=1}^{3} \sum_{\substack{P,Q \\ \left(\substack{P=I,J \\ Q=I,J}\right)}} [P^\alpha Q^\beta]_r \, \Delta\chi_p^\alpha \Delta\chi_q^\beta + \ldots \tag{A3}$$

$$\Delta\theta_{ijk} = \sum_{\alpha=1}^{3} I_\theta^\alpha \Delta\chi_i^\alpha + \sum_{\alpha=1}^{3} J_\theta^\alpha \Delta\chi_j^\alpha + \sum_{\alpha=1}^{3} K_\theta^\alpha \Delta\chi_k^\alpha + \tfrac{1}{2}\sum_{\alpha,\beta=1} \sum_{\substack{P,Q \\ \left(\substack{P=I,J,K \\ Q=I,J,K}\right)}} [P^\alpha Q^\beta]_\theta \, \Delta\chi_p^\alpha \Delta\chi_q^\alpha + \ldots \tag{A4}$$

$$\Delta\phi_{ijk\ell} = \sum_{\alpha=1}^{3} I_\phi^\alpha \Delta\chi_i^\alpha + \sum_{\alpha=1}^{3} J_\phi^\alpha \Delta\chi_j^\alpha + \sum_{\alpha=1}^{3} K_\phi^\alpha \Delta\chi_k^\alpha + \sum_{\alpha,\beta=1}^{3} L_\phi^\alpha \Delta\chi_\ell^\alpha$$
$$+ \tfrac{1}{2}\sum_{\alpha,\beta=1} \sum_{\substack{P,Q \\ \left(\substack{P=I,J,K,L \\ Q=I,J,K,L}\right)}} [P^\alpha Q^\beta]_\phi \, \Delta\chi_p^\alpha \Delta\chi_q^\beta + \ldots \tag{A5}$$

$$\Delta\delta_{ijkm} = \sum_{\alpha=1}^{3} I_\delta^\alpha \Delta X_i^\alpha + \sum_{\alpha=1}^{3} J_\delta^\alpha \Delta X_j^\alpha + \sum_{\alpha=1}^{3} K_\delta^\alpha \Delta X_k^\alpha + \sum_{\alpha=1}^{3} M_\delta^\alpha \Delta X_m^\alpha$$
$$+ \tfrac{1}{2} \sum_{\alpha,\beta=1}^{3} \sum_{\substack{P,Q \\ \left(\substack{P=I,J,K,M \\ Q=I,J,K,M}\right)}} [P^\alpha Q^\beta]_\delta \, \Delta X^\alpha \Delta X^\beta + \ldots \qquad \text{(A6)}$$

The coefficients in the transformations, I_r^α, J_r^α, I_θ^α, J_θ^α, K_θ^α ..., etc., are the derivatives

$$I_r^\alpha = \left.\frac{\partial r_{ij}}{\partial X_i^\alpha}\right|^o, \qquad J_r^\alpha = \frac{\partial r_{ij}}{\partial X_j^\alpha}, \qquad I_\theta^\alpha = \left.\frac{\partial \theta_{ijk}}{\partial X_i^\alpha}\right|^o$$

$$J_\theta^\alpha = \left.\frac{\partial \theta_{ijk}}{\partial X_j^\alpha}\right|^o, \qquad K_\theta^\alpha = \left.\frac{\partial \theta_{ijk}}{\partial X_k^\alpha}\right|^o, \qquad \text{etc.,}$$

and the coefficients $[I^\alpha I^\beta]_r$, $[I^\alpha J^\beta]_r$, ... $[I^\alpha I^\beta]_\theta$, $[I^\alpha I^\beta]_\theta$, etc., are the derivatives

$$[I^\alpha I^\beta]_r = \left.\frac{\partial^2 r_{ij}}{X_i^\alpha \; X_i^\beta}\right|^o, \qquad [I^\alpha J^\beta]_r = \frac{\partial^2 r_{ij}}{X_i^\alpha \; X_j^\beta},$$

$$[I^\alpha I^\beta]_\theta = \left.\frac{\partial^2 \theta_{ijk}}{X_i^\alpha \; X_i^\beta}\right|^o, \qquad [I^\alpha J^\beta]_\theta = \left.\frac{\partial^2 \theta_{ijk}}{X_i^\alpha \; X_j^\beta}\right|^o, \qquad \text{etc.}$$

$$I_\phi^{(\alpha)} = \frac{\partial \cos\phi_{ijk\ell}}{\partial X_i^\beta}, \qquad \text{etc.}$$

They may be calculated from the trial Cartesian coordinates, $\ldots\chi_i^{o(\alpha)}\ldots$, from the following considerations.

Since

$$r_{ij} = \left(\chi_j^{(1)} - \chi_i^{(2)}\right)^2 + \left(\chi_j^{(2)} - \chi_i^{(2)}\right)^2 + \left(\chi_j^{(3)} - \chi_i^{(3)}\right)^2 , \qquad \text{(A7)}$$

then,

$$I_r^\alpha = - \frac{\left(\chi_j^{o(\alpha)} - \chi_i^{o(\alpha)}\right)}{r_{ij}^o} , \qquad \text{(A8)}$$

$$J_r^\alpha = \frac{\left(\chi_j^{o(\alpha)} - \chi_i^{o(\alpha)}\right)}{r_{ij}^o} \qquad \text{(A9)}$$

and

$$[I^\alpha I^\beta]_r = [J^\alpha J^\beta]_r = -[I^\alpha J^\beta]_r = \frac{\delta^{\alpha\beta}}{r_{ij}^o} - \frac{\left(\chi^{o(\alpha)} - \chi^{o(\alpha)}\right)\left(\chi^{o(\beta)} - \chi^{o(\beta)}\right)}{{r_{ij}^o}^3} \qquad \text{(A10)}$$

where

$$\delta^{\alpha\beta} = \begin{cases} 1, \alpha=\beta \\ 0, \alpha\neq\beta \end{cases} .$$

In a similar manner, I_θ^α, J_θ^α, and K_θ^α can be calculated from

$$\cos\theta_{ijk} = - \frac{\underset{\sim}{r}_{jk} \cdot \underset{\sim}{r}_{ij}}{r_{jk} r_{ij}} \qquad \text{(A11)}$$

and

$$d\theta_{ijk} = \frac{1}{r_{jk}^o r_{ij}^o \sin\theta_{ijk}^o} \left[\underset{\sim}{r}_{jk}^o \cdot d\underset{\sim}{r}_{ij} + d\underset{\sim}{r}_{jk} \cdot \underset{\sim}{r}_{ij}^o - \frac{(\underset{\sim}{r}_{jk} \cdot \underset{\sim}{r}_{ij})(r_{ij} dr_{jk} - r_{jk} dr_{ij})}{r_{ij}^o r_{jk}^o} \right] . \qquad \text{(A12)}$$

Collection of ΔX_i^α, ΔX_j^α, and ΔX_k^α coefficients in Eq. (A12) results in

$$I_\theta^\alpha = \frac{1}{\alpha_{ijk}}\left[-\left(X_k^{o(\alpha)} - X_j^{o(\alpha)}\right) + \frac{\beta}{r_{ij}^{o2}}\left(X_j^{o(\alpha)} - X_i^{o(\alpha)}\right)\right] \quad \text{(A13)}$$

$$J_\theta^\alpha = \frac{1}{\alpha_{ijk}}\left[-\left(1 + \frac{\beta}{r_{ij}^{o2}}\right)\left(X_j^{o(\alpha)} - X_i^{o(\alpha)}\right) + \left(1 + \frac{\beta}{r_{ik}^{o2}}\right)\left(X_k^{o(\alpha)} - X_j^{o(\alpha)}\right)\right] \quad \text{(A14)}$$

$$K_\theta^\alpha = \frac{1}{\alpha_{ijk}}\left[\left(X_j^{o(\alpha)} - X_i^{o(\alpha)}\right) - \frac{\beta}{r_{jk}^{o2}}\left(X_k^{o(\alpha)} - X_j^{o(\alpha)}\right)\right] \quad \text{(A15)}$$

where

$$\alpha_{ijk} = r_{ij}^o\, r_{jk}^o \sin\theta_{ijk}^o$$

$$\beta = r_{ij}\, r_{jk} \cos\theta_{ijk}^o$$

The second derivatives, $[I^\alpha I^\beta]_\theta$, ..., etc., could be calculated by differentiation of Eqs. (A13) to (A15). However, the result is sufficiently complicated that we have chosen to calculate these coefficients by numerical differentiation of Eq. (A11). For the ϕ transformation we use

$$\cos\phi_{ijk\ell} = \frac{(\underset{\sim}{r}_{jk}\times\underset{\sim}{r}_{ij})\cdot(\underset{\sim}{r}_{k\ell}\times\underset{\sim}{r}_{jk})}{r_{ij}r_{jk}\sin\theta_{ijk}\; r_{jk}r_{k\ell}\sin\theta_{jk\ell}} \quad \text{(A16)}$$

$$= \frac{(\underset{\sim}{r}_{k\ell}\cdot\underset{\sim}{r}_{jk})(\underset{\sim}{r}_{jk}\cdot\underset{\sim}{r}_{ij}) - (\underset{\sim}{r}_{k\ell}\cdot\underset{\sim}{r}_{ij})r_{jk}^2}{r_{ij}r_{jk}\sin\theta_{ijk}\; r_{jk}r_{k\ell}\sin\theta_{jk\ell}} \quad . \quad \text{(A17)}$$

Again, formulas for the transformation derivatives may be calculated from the transformation relation (A17), but we have chosen to calculate them by numerical differentiation, from

$$I_{\phi}^{\alpha} = \frac{\partial \cos\phi_{ijk\ell}}{\partial X^{\alpha}} \ldots, \qquad [I^{\alpha} I^{\beta}]_{\phi} = \frac{\partial^2 \cos\phi_{ijk\ell}}{\partial X_i^{\alpha}\, \partial X_j^{\beta}} \ldots, \text{ etc.} \tag{A18}$$

The angle δ_{ijkm} is calculated from the angle between the normal to the plane ijk and the vector from j to m as,

$$\cos(\pi/2 - \delta_{ijkm}) = \frac{\underset{\sim}{r}_{jk}\ (\underset{\sim}{r}_{jk} \times \underset{\sim}{r}_{ij})}{r_{jm} r_{jk} r_{ij}\ \sin\theta_{ijk}} \ . \tag{A19}$$

We shall be interested only in small displacements of m from coplanarity with $i,j,k(\delta_{ijkm}^{o} = 0)$, in which case differentiation of Eq. (A23) and collection of coefficients X_i^{α}, X_j^{α}, X_k^{α}, and X_m^{α} result in

$$I_{\delta}^{\alpha} = \frac{(\underset{\sim}{r}_{jk}^{o} \times \underset{\sim}{r}_{jm}^{o}) \cdot \underset{\sim}{u}^{\alpha}}{\underset{\sim}{r}_{jm}^{o}\ \alpha_{ijk}} \tag{A20}$$

$$J_{\delta}^{\alpha} = \frac{(\underset{\sim}{r}_{jm}^{o} \times \underset{\sim}{r}_{ik}^{o}) \cdot \underset{\sim}{u}^{\alpha} - (\underset{\sim}{r}_{jk}^{o} \times \underset{\sim}{r}_{ij}^{o}) \cdot \underset{\sim}{u}^{\alpha}}{r_{jm}^{o}\ \alpha_{ijk}} \tag{A21}$$

$$K_{\delta}^{\alpha} = \frac{(\underset{\sim}{r}_{ij}^{o} \times \underset{\sim}{r}_{jm}^{o}) \cdot \underset{\sim}{u}^{\alpha}}{r_{jm}\ \alpha_{ijk}} \tag{A22}$$

$$M_{\delta}^{\alpha} = \frac{[(\underset{\sim}{r}_{jk} \times \underset{\sim}{r}_{ij}) \cdot \underset{\sim}{u}^{\alpha}]}{r_{jm}^{o}\ \alpha_{ijk}} \tag{A23}$$

where $\underset{\sim}{u}^\alpha$ is a unit vector along the χ^α axis.

Substitution of Eqs. (A3) to (A6) into Eq. (A2) results in the following equation for the potential energy

$$
\begin{aligned}
U = {} & U^o(\chi_i^{o(\alpha)} \ldots \chi_N^{o(\alpha)}) + \sum_{(ij)} U_r'(ij)\Bigg[\sum_{\alpha=1}^{3}(I_r^\alpha\Delta\chi_i^\alpha + J_r^\alpha\Delta\chi_j^\alpha) \\
& + \tfrac{1}{2}\sum_{\alpha,\beta=1}\sum_{\substack{P,Q\\(P=I,J)\\(Q=I,J)}}[P^\alpha Q^\beta]_r\,\Delta\chi^\alpha\Delta\chi^\beta\Bigg] + \tfrac{1}{2}\sum_{(ij)}U_r''(ij)\Bigg[\sum_{\alpha=1}^{3}(I_r^\alpha\Delta\chi_i^\alpha + J_r^\alpha\Delta\chi_j^\alpha)\Bigg]^2 \\
& + \sum_{(ijk)}U_\theta'(ijk)\Bigg[\sum_{\alpha=1}^{3}(I_\theta^\alpha\Delta\,{}_i^\alpha + J_\theta^\alpha\Delta\,{}_j^\alpha + K_\theta^\alpha\Delta\,{}_k^\alpha) + \tfrac{1}{2}\sum_{\alpha,\beta=1}\sum_{\substack{P,Q\\(P=I,J,K)\\(Q=I,J,K)}}[P^\alpha Q^\beta]_\theta\,\Delta\chi_p^\alpha\Delta\chi_q^\beta\Bigg] \\
& + \tfrac{1}{2}\sum_{(ijk)}U_\theta''(ijk)\Bigg[\sum^{3}(I_\theta^\alpha\Delta\chi_i^\alpha + J_\theta^\alpha\Delta\chi_j^\alpha + K_\theta^\alpha\Delta\chi_k^\alpha)\Bigg]^2 + \sum_{(ijk\ell)}U_\phi'(ijk\ell) \\
& \times\Bigg[\sum_{\alpha=1}(I_\phi^\alpha\Delta\chi_i^\alpha + J_\phi^\alpha\Delta\chi_j^\alpha + K_\phi^\alpha\Delta\chi_k^\alpha + L_\phi^\alpha\Delta\chi_\ell^\alpha) + \tfrac{1}{2}\sum_{\alpha,\beta=1}^{3}\sum_{\substack{P,Q\\(P=I,J,K,L)\\(Q=I,J,K,L)}}[P^\alpha Q^\beta]_\phi\,\Delta\chi_p^\alpha\Delta\chi_q^\beta\Bigg] \\
& + \tfrac{1}{2}\sum_{(ijk\ell)}U_\phi''(ijk\ell)\Bigg[\sum_{\alpha=I}^{3}(I_\phi^\alpha\Delta\chi_i^\alpha + J_\phi^\alpha\Delta\chi_j^\alpha + K_\phi^\alpha\Delta\chi_k^\alpha + L_\phi^\alpha\Delta\chi_\ell^\alpha)\Bigg]^2 \\
& + \sum_{(ijk)}U_\delta'(ijk)\Bigg[\sum_{\alpha=1}^{3}(I_\delta^\alpha\Delta\chi_i^\alpha + J_\delta^\alpha\Delta\chi_j^\alpha + K_\delta^\alpha\Delta\chi_k^\alpha + M_\delta^\alpha\Delta\chi_m^\alpha) \\
& + \tfrac{1}{2}\sum_{\alpha,\beta=1}\sum_{\substack{P,Q\\(P=I,J,K,M)\\(Q=I,J,J,M)}}[P^\alpha Q^\beta]_\delta\,\Delta\chi_p^\alpha\Delta\chi_q^\beta\Bigg] + \tfrac{1}{2}\sum_{(ijkm)}U_\delta''(ijkm) \\
& \times\Bigg[\sum_{\alpha=1}^{3}(I_\delta^\alpha\Delta\chi_i^\alpha + J^\alpha\Delta\chi_j^\alpha + K_\delta^\alpha\Delta\chi_k^\alpha + M_\delta^\alpha\Delta\chi_m^\alpha)\Bigg]^2
\end{aligned}
\qquad (A24)
$$

where

$$U_r'(ij) = \left.\frac{\partial U}{\partial r_{ij}}\right|_{r_{ij}=r_{ij}^o} ,$$

$$U_r''(ij) = \left.\frac{\partial^2 U}{\partial r_{ij}^2}\right|_{r_{ij}=r_{ij}^o} , \quad \text{etc.} \tag{A25}$$

Application of the necessary condition for a minimum in U,

$$\frac{\partial U}{\partial X_i^\alpha} = 0 \begin{pmatrix} \alpha=1,2,3 \\ i=1,\ldots N \end{pmatrix} \tag{A26}$$

to Eq. (A24) leads to a set of linear algebraic equations for the ΔX^α, which may be solved by standard methods. The ΔX^α values determined lead to a new conformation which minimizes (or maximizes) U in Eq. (A24). This will not, in general, minimize U in Eq. (A1), since the expansion of the potential in Eq. (A2) and the expansions in the transformations, Eqs. (A3) to (A6), are approximate. However, the ΔX^α values may be used to calculate a new set of trial coordinates,

$$X_i^{o(\alpha)}(\text{new}) = X_i^{o(\alpha)}(\text{old}) + \Delta X_i^\alpha , \tag{A27}$$

a new set of derivatives, Eq. (A25) and coefficients, Eqs. (A8, 9, 10, 13) to (15, 18,20) to Eq. (A23), and the minimization repeated. When, after repeated iteration, the ΔX^α are zero to within prescribed limits, the iteration can be terminated and the geometrical factors of interest calculated from the final set of trial coordinates.

In the final iteration the coefficient of each ΔX^{α} in the *linear* terms in Eq. (A24) are zero (very nearly) and only the quadratic terms remain. This final potential may then be used to calculate the vibrational frequencies of the molecule by standard methods. In turn, these frequencies may be used to calculate the vibrational heat capacity and other thermodynamic functions.

APPENDIX B

(Taken from the Ph.D. Dissertation of D.H. Wertz, submitted to the University of Georgia, 1974)

NEWTON RAPHSON

In the Newton-Raphson method one assumes that the potential energy surface can be approximated by a Taylor's Series terminated after the second order terms:

$$V = V^o + \sum_{i=1} \Delta\xi_i \left(\frac{\partial V}{\partial \xi_i}\right)_o + \tfrac{1}{2}\sum_{i=1}^{3N}\sum_{j=1}^{3N} \Delta\xi_i \Delta\xi_j \left(\frac{\partial V}{\partial \xi_i \partial \xi_j}\right)_o \qquad (1)$$

where ξ_i = the coordinates of the molecule, and V = steric energy.

The Taylor's Series above implies that there is an equation like Eq. (2) that is a good approximation to the potential energy surface about the present set of coordinates,

$$V = V^o + \sum_{i=1}^{3N} A_i \xi_i + \sum_{i=1}^{3N}\sum_{j=1}^{3N} B_{ij}\ \xi_i \xi_j \qquad (2)$$

where A_i and B_{ij} are constants.

The equations above give V at a minimum when all of the partial derivatives are equal to zero. The partial derivatives can also be approximated by a Taylor's Series that is terminated after the linear terms.

$$\frac{\partial V}{\partial \xi_i} = \frac{\partial V^o}{\partial \xi_i} + \sum_{j=1}^{3N} \Delta\xi_j \frac{\partial\left(\frac{\partial V}{\partial \xi_i}\right)}{\partial \xi_j} \quad , \qquad \text{for } i = 1,2,\ldots,3N \ . \qquad (3)$$

We are interested in the set of $\Delta\xi_j$ such that

$$0 = \frac{\partial V^o}{\partial \xi_i} + \sum_{j=1}^{3N} \Delta\xi_j \frac{\partial^2 V}{\partial \xi_i \partial \xi_j} , \qquad \text{for } i = 1,2,\ldots,3N .$$

The above set of linear equations cannot be solved as is, because as is well known, there are only 3N-6 internal degrees of freedom in a molecule while there are 3N equations above. This means there is more than one set of $\Delta\xi_i$ which will solve the equations. In order that a nontrivial solution be found, it is necessary to remove the six translational-rotational terms. Boyd[54] does this by fixing six of the coordinates of the molecule such that the molecule cannot translate or rotate. This means all of the derivatives with respect to these coordinates can be removed from the matrix. Lifson[51] gets around the problem by expressing the energy in terms of normal coordinates instead of Cartesian coordinates.

The assumption that the Taylor's Series can be terminated after the quadratic term is, of course, not exactly correct, so that several iterations of this procedure are necessary to minimize the energy of the molecule.

MODIFIED NEWTON-RAPHSON

Shortly after Boyd[54] published his Newton-Raphson scheme I started work on a minimization scheme that is a combination of the steepest descent method and a Newton-Raphson method where only the diagonal terms in the matrix are looked at.

This scheme is generally referred to in the research group as the *first derivative scheme.* The three partial derivatives with respect to the three coordinates of the atom are analytically calculated and the atom is then moved along each of the axes by an amount proportional to the derivatives. The three partial derivatives are then recalculated and the minimum energy position of the atom is calculated assuming that the derivative of the potential energy with respect to each of the coordinates can be approximated by

$$\frac{\partial V}{\partial \xi_i} = \frac{\partial V^o}{\partial \xi_i} + \Delta\xi_i \frac{\partial^2 V}{\partial \xi_i^2} \quad , \tag{5}$$

where ξ_i = X,Y,Z.

The second derivatives $\partial^2 V_2/\partial\xi_i^2$ are numerically calculated using the formula

$$\frac{\left(\frac{\partial V}{\partial \xi_i} - \frac{\partial V}{\partial \xi_i}\right)}{\left(\xi_i - \xi_i'\right)} = \frac{\partial^2 V}{\partial \xi_i^2} \tag{5a}$$

Assuming Eq. (5) is a good approximation to the derivatives of the energy of the molecule, with respect to the coordinates of the atom, is the same thing as assuming that the energy of the molecule can be approximated by

$$V = V^o + BX + CX^2 + DY + EY^2 + FZ + GZ^2 \quad . \tag{6}$$

The process of placing the atoms in the minimum energy position predicted by the equations above is done with each of the atoms in turn and the whole process is repeated until either the largest movement of any atom on any iteration falls below a fixed value or the energy after a few iterations (usually five) fails to decrease by a significant amount.

As is done in Wiberg's minimization scheme, atoms bonded to only one atom (mainly hydrogens) are moved with the atom they are bonded to, in addition to being looked at independently.

The only comparison made of the times required to minimize the first derivative scheme and Boyd's program was on n-hexane. In this particular case, the first derivative scheme was faster by about a factor of two. One would expect that the time required to do an iteration by Boyd's program would go up as the cube of the number of atoms in the molecule because the required time to solve a set of simultaneous linear equations goes up as the cube of the number of equations in the matrix.[50b] On the other hand, the time required per iteration for the first derivative scheme should be proportional to the number of interactions, which increases as the square of the number of atoms.

It is possible that further testing would have shown that these considerations were incorrect or that n-hexane was not representative of the average molecule, but the chances of this were not felt to be great enough to justify the effort necessary to find out. Boyd's program required so much core storage that it would have been difficult getting the University of Georgia Computer Center to give reasonable turnaround on jobs using this minimization scheme.

A somewhat more extensive comparison using four compounds was made between the steepest descent program and the first derivative scheme. On these four molecules it was found that the first derivative scheme was faster by a factor of 150. This was so much faster that further testing to determine which one was superior was felt to be unnecessary.

After the first derivative minimization scheme had been in operation for awhile, it became apparent that the minimization scheme could be improved if the potential energy surface was looked at as a quadratic surface with the cross terms included, and that the program could be speeded up if the second derivatives were calculated analytically (i.e., in one pass) rather than numerically. It seemed reasonable to assume that these changes would both cut the time required to do an iteration and would also enable the program to more accurately place the atom at its minimum position.

When these things were done the results were as hoped. Both the time per iteration and the number of iterations required to minimize the energy of the molecule decreased. The net improvement was about a factor of three.

The present minimization scheme essentially does a Newton-Raphson minimization on each atom in the molecule on each iteration. This is faster than doing a full Newton-Raphson on the complete molecule because one does not have to calculate the second partial derivatives to fill a $3N \times 3N$ matrix, and even more importantly, one does not have to diagonalize a large matrix on each iteration. In the present minimization scheme, an atom with hydrogen, or other atom which has only one bond bonded to it, has these atoms moved with it when the atom is moved. This means all interactions involving attached atoms must be looked at when one is looking at the atom to which they are bonded. The attached atoms are also looked at independently. This means interactions involving attached atoms are looked at twice which obviously increases the time required to do an iteration. However, this extra time is more than made up for by the fact that the atoms move more rapidly to their minimum position.

PARTITIONING OF ONE-PARTICLE SPACE

- Internal space: the subspace spanned by the MCSCF occupied orbitals
- External space: the orthogonal complement of the internal space
- Inactive space: the subspace spanned by the occupied orbitals frozen in the CI calculation
- Active space: the subspace spanned by occupied orbitals excited in the CI calculation
- Classify configurations by (i,e), where i is the number of electrons occupying internal orbitals, e is the number of electrons occupying external orbitals

COMPLETE ACTIVE ELECTRON CI

- All possible configurations with n electrons in the active orbitals (n,0)
- Proper dissociation
- Near-degeneracy effects
- Size-consistency
- Qualitative PES for reactions (PES ≡ Potential Energy Surface)
- Good for bond lengths, usually too long
- Good for frequencies, usually too small
- Poor for dissociation energies, usually too small

- Poor for excitation energies, sometimes wrong order, but good in some cases

SINGLY EXCITED CI, FIRST ORDER CI, POL-CI, OVC, etc.

- (n,0) + (n-1,1) configuration
- Nearly correct dissociation
- Polarization and semi-internal correlation effects
- Based on the qualitative idea of separation of atomic and molecular correlation energies
- Slightly better than (n,0) for bond lengths and frequencies
- Much better than (n,0) for dissociation and excitation energies

DOUBLY EXCITED CI

- (n,0) + (n-1,1) + (n-2,2) configurations
- Not size-consistent
- About the biggest CI we can do
- Good dissociation energies?
- Good ionization potentials?
- Potential surfaces for reactions 2-3 kcal/mole accuracy for barriers?

- All three types of calculations are independent of choice of virtual orbitals
- These calculations can get very large and we need ways of selecting important configurations

CONFIGURATION SELECTION, PERTURBATION THEORY

$$H = H_o + \lambda V$$

$$\Psi = \sum_n \lambda^n \psi^n$$

$$E = \sum_n \lambda^n E^n$$

$$H_o \psi^o = E^o \psi^o$$

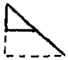
$$H_o \psi^1 + V\psi^o = 0$$

$$H_o \psi_k^o = E_k^o \psi_k^o$$

$$E_o^2 = -\sum_k \frac{\langle \psi_k^o | H | \psi_o^o \rangle^2}{\langle \psi_k^o | H | \psi_k^o \rangle - E_o^o}$$

In the CI context, choose $\{\Phi_I^o\}$ and make

$$H_o = \{\Phi_I^o\} \quad \{$$



Classify configurations by the order of perturbation wavefunction in which they first appear.

INTERACTING SUBSPACES

- Zeroth order subspace $\{\Phi_I^o\}$ chosen by chemical or energetic considerations
- First order interacting subspace $\{\Phi_J^1\}$
  $\langle\Phi_J^1|H|\Phi_I^o\rangle \neq 0$ for some I
- Second order interacting subspace $\{\Phi_k^2\}$
  $\langle\Phi_k^2|H|\Phi_J^1\rangle \neq 0$ for some J
- One may, for example, keep only configurations in the first order interacting subspace with respect to a chosen zeroth order subspace
- The configurations included in the MCSCF-calculation is often a good choice for the zeroth order space
- The resulting CI is still independent of the choice of virtual orbitals

FURTHER SELECTION OF CONFIGURATIONS

- Energy selection: discard all configurations whose estimated contribution is below some threshold.

  i) perturbation theory

$$\Delta E_k = \frac{\langle\Phi_k|H|\Phi^o\rangle^2}{(H_{kk} - E^o)}$$

  ii) $\Phi^k = (1 - C_k^2)^{-\frac{1}{2}} \sum_{I\neq k} C_I \Phi_I$

$$\Delta E_k = (E - H_{kk})C_k^2 / (1 - C_k^2)$$

ΔE and C_k estimated by BK method

Advantage: significant reduction.

Disadvantage: bumpy energy surfaces, properties not as good as energy, and depends on virtual orbitals.

- Natural orbital truncation: discard all natural orbitals with occupation numbers below some threshold and carry out the CI calculation in the reduced basis set.

  Advantages: smooth potential surface, good properties, good for excitation energies, and useful for treating higher order subspaces.

  Disadvantages: not efficient for total energy or dissociation energy, costly to obtain but approximate NO from perturbation and wavefunction seems to work well.

EXTRAPOLATION BASED ON ENERGY SELECTION

Calculate CI energies for different values of the threshold. Extrapolate to the zero threshold limit.

Advantage: can get close estimate of CI energy by a series of relatively small calculations.

Disadvantages: reliable properties? extrapolation procedures not always reliable, bumpy surface, and virtual orbital-dependent.

CONSTRUCTION OF ORBITAL BASIS $\{\phi_i\}$

Internal orbitals:

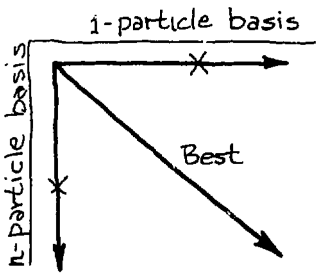
1) MCSCF (including GVB)
2) IVO, ICVO, etc., virtual orbitals determined in n-1 potential, good for singly excited states

Virtual orbitals:

1) Approximate natural orbitals from perturbation wavefunctions, best for energy selection and orbital truncation
2) IVO, ICVO, etc., not very useful for energy selection
3) Virtual orbitals determined in n-V potential, where V is the number of valence electrons -- good for energy selection

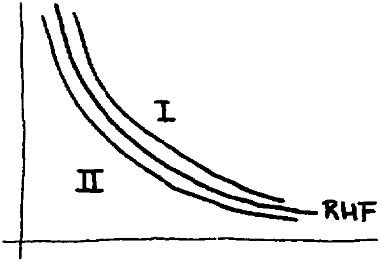
HOW TO CHOOSE BASIS SETS?

The basis set error should be a small part of the remaining error in the calculated properties. Balance between n-particle basis set and 1-particle basis set



$\underline{He_2 \text{ GROUND STATE:}}$ $1\sigma_g^2\, 1\sigma_u^2$

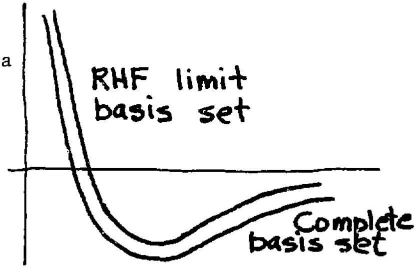
- RHF potential curve is known
- How do we find a basis set that gives an SCF curve parallel to the RHF limit curve?
- Calculation I:
  complete atomic basis,
  no polarization functions,
  SCF curve too repulsive,
  need p and d functions
  to describe distortion
  of atoms.



- *Calculation II:*
  limited atomic basis set, say double-zeta with polarization functions. SCF curve not repulsive enough because of *basis set superposition error*, basis functions on one center are helping improve the description of the other atom, leading to an artificial lowering.

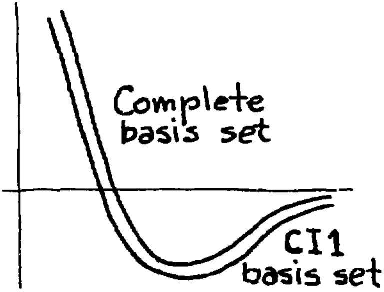
$\underline{He_2 \text{ CI}}$

- CI1 $1s_a^2\, 1s_b^2 + \uparrow a + \uparrow b + \uparrow a \uparrow b$
  ($\uparrow a \equiv$ singly excited configurations from orbitals on center a)
- RHF limit basis set yields a potential curve that is not attractive enough. Needs more diffuse polarization functions.



- CI2 $(1s_a^2 + 2s_a^2 + 2p_a^2)(1s_b^2 + 2s_b^2 + 2p_b^2) + \dagger a + \dagger b + \dagger a \dagger b$

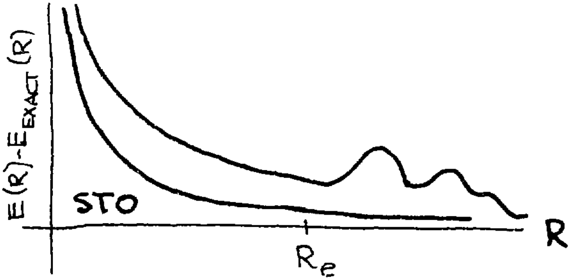
- CI1 basis set yields a potential curve that is too attractive, because superposition error is back. To describe the additional atomic correlation, additional atomic basis functions are needed.



$\underline{H_2}$

- Complete CI using 4s-3p-1d (STO) and 4s-3p-1d (CGTO) basis sets

- Comp with exact results



- The bumps in the error function for the CGTO calculation is caused by the incorrect long-range behavior of GTO -- they die off too quickly.

Energy difference at R_e between CH $a\,^4\Sigma^-$ and $A\,^2\Delta$

Natural orbital truncation, all singles and doubles from HF

σ	π	$E(^4\Sigma^-)$	Error	$E(^2\Delta)-E(^4\Sigma)$	Error
4	1	-38.304738	0.014373	0.107820	0.000948
6	-	-38.314663	0.004448	0.107936	0.000832
8	-	-38.317600	0.001511	0.108720	0.000048
23	-	-38.319111	-	0.108768	-
23	3	-38.356300	0.007666	0.099552	0.002149
-	6	-38.362760	0.001206	0.097735	0.000364
-	8	-38.363515	0.000451	0.097574	0.000203
-	13	-38.363966	-	0.097371	-

- Convergence on energy difference much better than convergence on total energy.

Comparison of Calculated and Experimental Spectroscopic Constants for B_2 ($X^3\Sigma_g^-$)

	$\Delta G_{1/2}$ (cm^{-1})	$\Delta G_{3/2}$ (cm^{-1})	R_e (Å)	D_e (eV)	($2^3\Sigma_u^-$)
Experiment	1026	1006	1.590		3.79
Δ(VCI)	+ 56	+ 51	+0.026	2.58	+0.85
Δ(FOCI)	+ 53	+ 49	+0.030	2.86	+0.11

VCI = Complete Active Electron CI.

FOCI = First Order CI.

- VCI as good as FOCI for bond lengths and frequencies
- FOCI more accurate for dissociation and excitation energies.

Comparison of Calculated and Experimental Spectroscopic Constants for O_2 ($X^3\Sigma_g^-$)

	$\Delta G_{1/2}$ (cm^{-1})	$\Delta G_{3/2}$ (cm^{-1})	R_e (Å)	D_e (eV)
Experiment	1568.5	1544.6	1.208	5.213
Δ(VCI)	- 97	- 98	+0.02	-1.43
Δ(FOCI)	- 87	- 84	+0.03	-0.26

- Same conclusions as before for VCI and FOCI.

Comparison of Calculated and Experimental Spectroscopic Constants for CH ($X^2\Pi$)

	$\Delta G_{1/2}(cm^{-1})$	$\Delta G_{3/2}(cm^{-1})$	$R_e(a_0)$	$D_e(eV)$	$T_e(a^4\Sigma^-)(eV)$
Observed	2732.5	2606.5	1.120	3.63	0.70
Δ(VCI)	-172.5	-185.7	+0.017	-0.68	-0.51
Δ(FOCI)	-161.5	-169.3	+0.020	-0.46	-0.11
Δ(SDHF)	- 10.4	- 48.5	-0.002	-0.12	-0.05

VCI and FOCI = same as before;

SDHF = single and double excitation from HF configuration.

- In this case, SDHF more accurate than VCI and FOCI.

Vibrational quanta of $CO(X^1\Sigma^4)$ in cm^{-1}.

Method	$\Delta G_{1/2}$	Error	$\Delta G_{3/2}$	Error
VCI	2128.3	-15.0	2102.4	-14.4
FOCI	2140.3	- 3.0	2114.7	- 2.1
SDHF	2235.9	92.6	2213.9	96.1
observed	2143.3	--	2116.8	--

- VCI and FOCI give much better vibrational quanta in spite of higher total energies.
- Needs six-fold excitations to describe the stretch of a triple bond.

Comparison of Calculated and Empirical Dipole Moment Function* for CO($X^1\Sigma^+$):

$$\mu(R) = \sum_{i=0}^{3} M_i (R - R_e)^i$$

	M_0	M_1	M_2	M_3
Empirical	-0.1222	1.645	-0.042	-0.350
Δ(VCI)	-0.138	+0.081	+0.041	-0.158
Δ(FOCI)	-0.197	-0.067	0	-0.026
Δ(SDHF)	+0.040	+0.225	+0.134	+0.829

*R in Å, μ in debye.

K.Kirby-Docken and B.Liu: Molecular dipole moment functions. 1

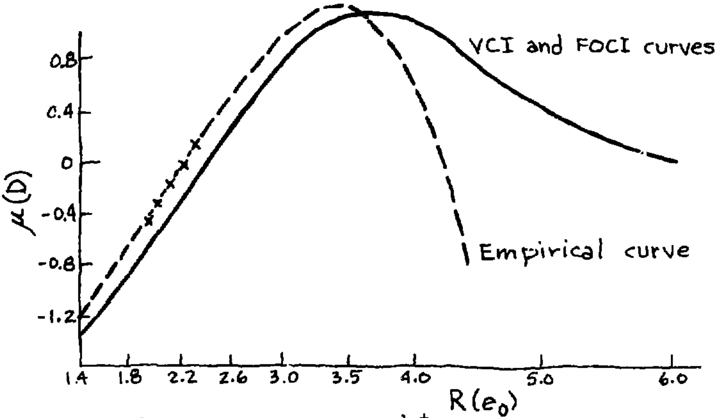


Fig. 1. Dipole moment functions for $X^1\Sigma^+$ state of CO. The solid curve goes through the FOCI points presented in Table V. The dashed curve is empirical dipole moment function of Young and Eachus.[10] The x's are points computed by Billingsley and Krauss.[18] The VCI results discussed in the text closely parallel the FOCI curve and would not be easily distinguishable in this figure.

- FOCI curve and empirical curve agree well for R ≤ 3.5 a.u.

Dipole moment of vibrational states of $CO(d^3\Delta)$

v	μ_{calc} (debye)		μ_{expt} (debye)	
3	-0.624		-0.48±0.02	
		0.060		0.06±0.04
4	-0.564		-0.42±0.02	
5				
		0.121		0.14±0.04
6	-0.433		-0.28±0.02	

- Theory-predicted results later confirmed by experiment.

Force constants of HCN in indyne/Å.

	K_1	K_2	K_3
Experiment	3.12(0.01)	9.39(0.09)	-0.21(0.08)
Δ(VCI)	-0.07	-0.15	0
Δ(SDHF)	+0.15	+0.94	+0.06

K_1 = CH stretch

K_2 = CN stretch

K_3 = coupling between CH and CN stretches

- VCI results better than SDHF results.
  6-fold excitation needed for stretching triple bond.

$\underline{O_2 \text{ STUDY}}$

R.P.Saxon
B.Liu

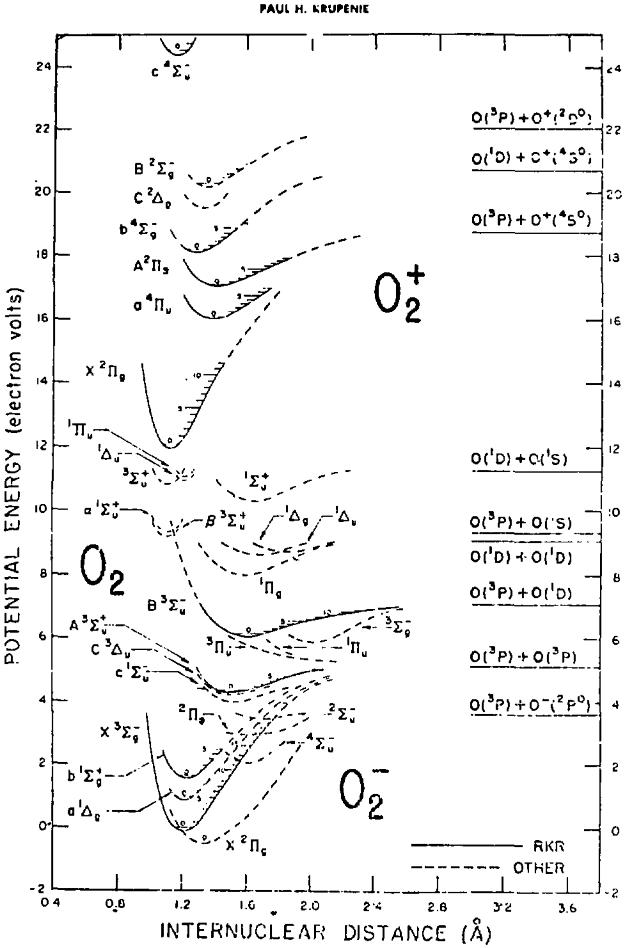
- FOCI calculations
- 62 valence states arising from oxygen ^3P, ^1D and ^1S
- The maximum error in calculated spectroscopic constants, for seven low-lying bound states are:

R_e – 0.04 Å

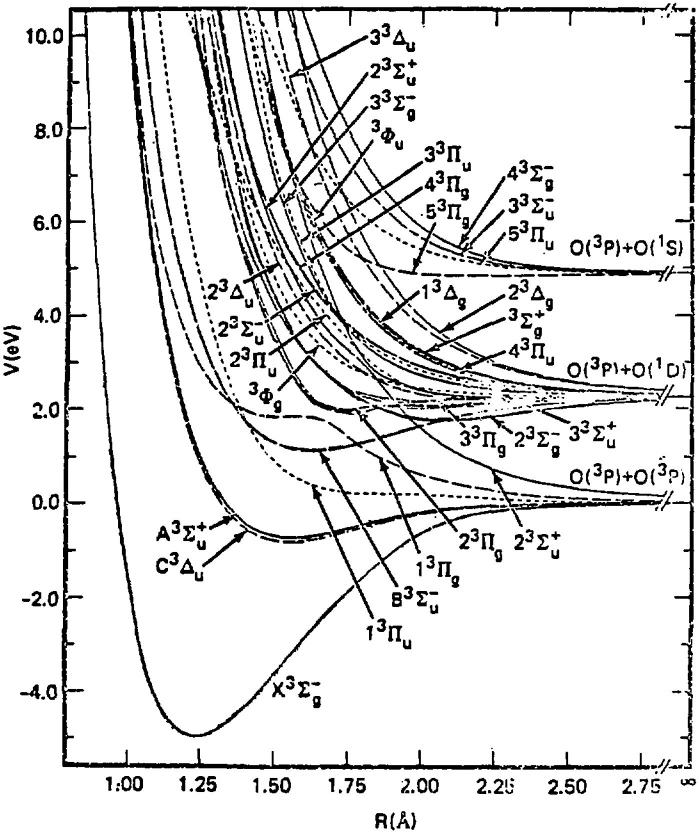
D_e – 0.4 eV

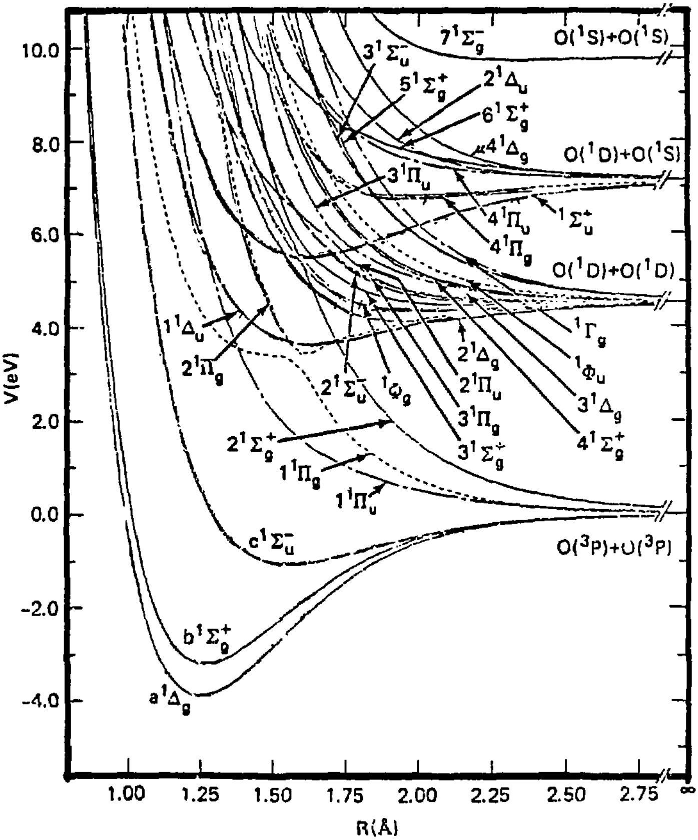
T_e – 0.2 eV

ω_e – 120 cm^{-1}

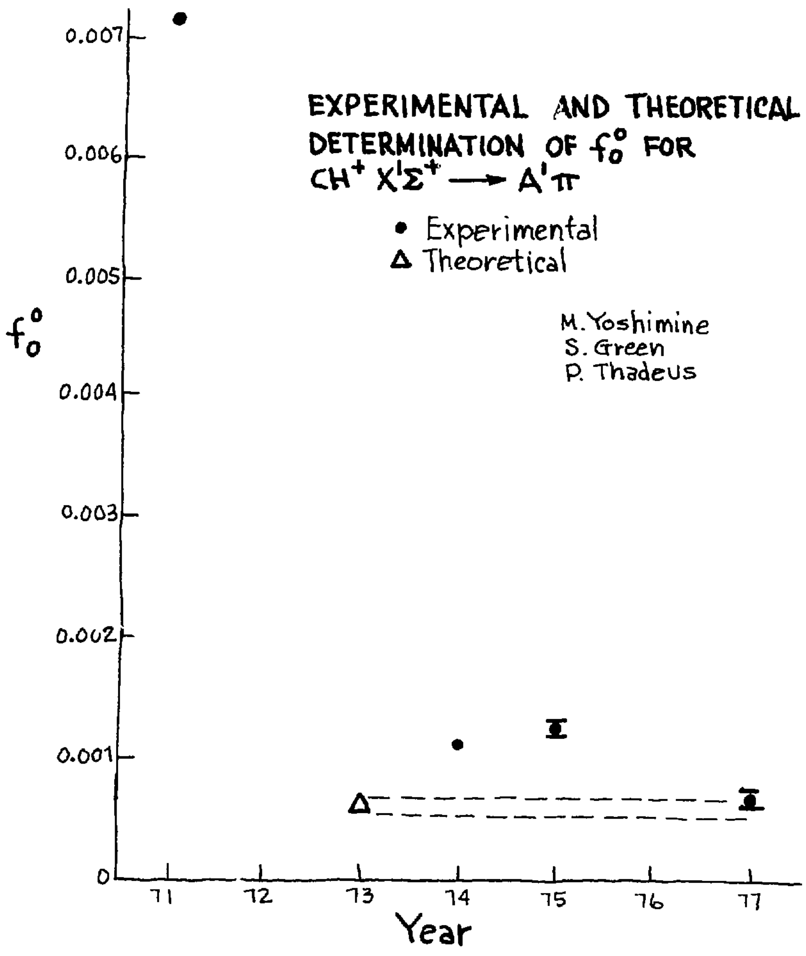


Potential energy curves for O_2, O_2^- and O_2^+.


V(eV)
R(Å)
10.0
8.0
6.0
4.0
2.0
0.0
-2.0
-4.0
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
∞
$3^3\Delta_u$
$2^3\Sigma_u^+$
$3^3\Sigma_g^-$
$^3\Phi_u$
$3^3\Pi_u$
$4^3\Pi_g$
$5^3\Pi_g$
$4^3\Sigma_g^-$
$3^3\Sigma_u^-$
$5^3\Pi_u$
O(^{3}P)+O(^{1}S)
$2^3\Delta_u$
$2^3\Sigma_u^-$
$2^3\Pi_u$
$^3\Phi_g$
$1^3\Delta_g$
$2^3\Delta_g$
$^3\Sigma_g^+$
$4^3\Pi_u$
O(^{3}P)+O(^{1}D)
$3^3\Pi_g$
$2^3\Sigma_g^-$
$3^3\Sigma_u^+$
O(^{3}P)+O(^{3}P)
$A^3\Sigma_u^+$
$C^3\Delta_u$
$2^3\Pi_g$
$2^3\Sigma_u^+$
$1^3\Pi_g$
$B^3\Sigma_u^-$
$1^3\Pi_u$
$X^3\Sigma_g^-$



7¹Σg⁻
O(¹S)+O(¹S)
3¹Σu⁻
5¹Σg⁺
2¹Δu
6¹Σg⁺
4¹Δg
O(¹D)+O(¹S)
3¹Πu
4¹Πu
¹Σu⁺
4¹Πg
O(¹D)+O(¹D)
1¹Δu
2¹Πg
¹Γg
2¹Δg
¹Φu
2¹Σu⁻
¹Φg
2¹Πu
3¹Δg
3¹Πg
2¹Σg⁺
3¹Σg⁺
4¹Σg⁺
1¹Πg
1¹Πu
c¹Σu⁻
O(³P)+O(³P)
b¹Σg⁺
a¹Δg
V(eV)
10.0
8.0
6.0
4.0
2.0
0.0
-2.0
-4.0
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
∞
R(Å)


EXPERIMENTAL AND THEORETICAL DETERMINATION OF f_0^o FOR CH^+ $X^1\Sigma^+ \longrightarrow A^1\Pi$
• Experimental
Δ Theoretical
M. Yoshimine
S. Green
P. Thadeus
f_0^o
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
71
72
73
74
75
76
77
Year

COMPUTATIONAL QUANTUM CHEMISTRY:
FUTURE OUTLOOK

Lecture 19

by

Professor Ernest R. Davidson
Department of Chemistry
University of Washington
Seattle, Washington

FUTURE OUTLOOK

NEW METHODS

Ab Initio Spin-Spin and Spin-Orbit

Perturbation Methods

Rayleigh-Schrodinger

$$\psi = \psi^{(0)} + \psi^{(1)} + \psi^{(2)} + \ldots$$

$$E = E^{(0)} + E^{(1)} + E^{(2)} + \ldots$$

$$\psi^{(n)} = \sum_i{}' C_i^{(n)} \psi_i^{(0)} \quad , \qquad n \neq 0$$

$$E = \langle\psi^o|H|\psi^o\rangle + \sum_{ij}{}' \langle\psi^o|H|\psi_i^o\rangle (E-\mathcal{H})_{ij}^{-1} \langle\psi_j^o|H|\psi^o\rangle$$

$$\mathcal{H}_{k\ell} = \langle\psi^o|H|\psi_\ell^o\rangle \quad , \quad \text{omit } \psi^{(0)} \text{ from set}$$

$$\mathcal{H} = \mathcal{H}^0 + \mathcal{H}^{(1)} \quad , \qquad E = E_o + \Delta$$

$$(E-\mathcal{H})^{-1} = \sum_{k=0}^{\infty} (E_o - \mathcal{H}^o)^{-1} \left[(\mathcal{H}^{(1)} - \Delta)(E_o - \mathcal{H}^o)^{-1}\right]^k$$

can be used to fourth order with reasonable MC-SCF ψ^0 to give reliable dissociation energy, properties, ionization energy, etc.

MBPT simplified version

$$\psi^o = \psi_{SCF}$$

$$H^o = \sum F(i)$$

less general, but simpler to compute.

Møller-Plesset Formula

$$E^o + E^1 = E_{UHF}$$

$$E^2 = -\sum_{i<j}^{occ} \sum_{a<b}^{virt} \frac{\langle ij|r_{12}^{-1}|ab-ba\rangle}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}$$

ε = orbital energy of canonical UHF orbitals

Third and fourth orders are more complicated but similar.

ANALYTICAL DERIVATIVES

$$\frac{\partial E}{\partial X_i} , \quad \frac{\partial^2 E}{\partial X_j \partial X_i}$$

Better search procedures for stationary points.

Continuum problems: absorption and scattering.

Dyson equations: time-dependent perturbation theory and Green's function.

EXTENDED RANGE OF PROBLEMS

1. Spectroscopy

Magnetic

$\underset{\approx}{A}$, $\underset{\approx}{D}$, $\underset{\approx}{Q}$ hyperfine, zero-field field-gradient tensors

Electric

$\underset{\sim}{\mu}$, $\frac{\partial \mu}{\partial x}$ dipole

$\underset{\sim}{\alpha}$, $\frac{\partial \alpha}{\partial x}$ polarizability tensor

- Photoionization: peak location and cross section.
- Electron excitation: Rydberg-valence mixing.

2. Structure
   - Bond lengths and angles
   - Multiple structures (rotomers)
   - Thermochemistry (ΔH, ΔS)

3. Kinetics
   - Isotope effects
   - Transition state, force constants, and structure
   - Spin-orbit rates
   - Reaction mechanisms
   - Surface sites and reactions
   - Photochemistry

What will remain hard?

- Condensed phase!

  Solvent effects of spectra, structure, kinetics.
- Secondary, tertiary, etc. structures of biological systems.
- High precision results for moderate-size molecules, i.e., singlet-triplet splits to ±2 kcal.
- Mixed valence excited states:

  MC-SCF with non-orthogonal orbitals
- Photochemistry:

  Jahn-Teller effects from surface crossings.

DEEPER UNDERSTANDING OF RESULTS

- What is objective of measurement of physical properties ($\vec{\mu}$, $\underset{\approx}{\alpha}$, $\underset{\sim}{A}$, $\underset{\sim}{D}_i$, etc.)? to determine electronic structure!

- What is objective of electronic spectroscopy?
  Assignment of energies to electronic structures!

- What is objective of chemical structure determination?
  To understand electronic structure!

- What is objective of kinetics?
  Predict rate and products of a reacting mixture

- What is (present) objective of biochemistry?
  To understand relation between structure and function.

- What is the purpose of quantum chemistry?
  To understand electronic structure!
  a. level of accuracy and relation to accurate predictions
  b. relation of structure to geometry
  c. relation of structure to reactivity

i.e., we start with what everyone else wanted to know.

Our problem is to relate our results to trends in chemical/physical properties.

Ab initio calculations on individual molecules cannot solve this problem. We need results for families of molecules.

Too much of quantum chemistry simply asks, "Here is an experimental result, can we reproduce it?" The answer is only interesting if it is, "No, because the experiment is wrong."

What do we know when the calculation is over that we did not know at the start (provided we agree with experiment)? The detailed electron distribution. The details of potential surface far from equilibrium geometry.

What results are usually reported in publications?

GPO 694-194/98