

# Northeast Regional Molecular Modeling Workshop 1998

University of Massachusetts at Amherst  
Chemistry Resource Center  
Goessmann Laboratory 152

## **Dr. Beatrice Botch**

NERMMW Program Administrator, UMass-Amherst  
Director, Chemistry Resource Center  
413-545-5257  
bbotch@chem.umass.edu

## **Prof. Paul M. Lahti**

NERMMW Instructor, UMass-Amherst  
413-545-4890  
lahti@chem.umass.edu

## **Prof. Vernon G. S. Box**

NERMMW Presenter, City College of New York  
exorga@csi.com

## **Dr. Robert Lancashire**

NERMMW Presenter, University of the West Indies-Mona, Jamaica  
rjlanc@uwimona.edu.jm

## **Prof. Eric Martz**

NERMMW Presenter, UMass-Amherst  
emartz@microbio.umass.edu

<http://www.chem.umass.edu/~nermmw>  
[nermmw@copper.chem.umass.edu](mailto:nermmw@copper.chem.umass.edu)



This project was supported, in part,  
by the

**National Science Foundation**

Opinions expressed are those of the authors  
and not necessarily those of the Foundation

## Northeast Regional Molecular Modeling Workshop

### *Staff*

Mr. Frank Stolle, CRC Network Administration

Mr. John Husband, Graduate Assistant

Ms. Jana Sorge, Administrative Assistant

### *Supporters and Industrial Participants*

UMass-Amherst

National Science Foundation (Division of Undergraduate Education)  
Grant DUE 9554634

Cambridgesoft, Inc.

Gaussian Inc.

Interactive Simulations, Inc.

Molecular Simulations, Inc.

Thomson Scientific Instrument Company

Wavefunction, Inc.

To the NSF and to those in industry who have supported us and participated with us, we  
are deeply grateful!

## NERMMW Summary

The NERMMWs are workshops on computational chemistry. They have been funded by the federal government with the purpose of training college faculty in the Northeastern USA to be able to use computational chemistry methodology in an educational setting. The grants supporting the NERMMWs allow participants to attend the workshops for free during the 3 1/2 days of the curriculum, with only travel costs being borne by the attendees. Educators from resource-poor institutions are provided a stipend of \$150. Women and minority candidates are encouraged to apply.

The workshops are kept small (10-15 participants) in order to maximize interaction with the instructors. The curriculum is intensive -- with morning, afternoon, and evening sessions -- but friendly and conversational. Various sessions are comprised of lectures concerning background, structured laboratory computations that demonstrate lecture precepts, and free-format periods with a variety of suggested activities that aim to let participants move at their own paces and try any of a variety of types of software and hardware. Hands-on instruction is maximized on a variety of modern computational hardware available through the UMass-Amherst Chemistry Department Chemistry Resource Center -- hardware used includes Pentium class PCs, Power PCs, and Silicon Graphics Indy workstations. All participants are able to work at their own computers, but in a shared laboratory atmosphere intended to encourage everyone to learn from the questions of each. Very little previous experience is expected for participants -- a working knowledge of Windows or Mac-OS is sufficient to allow one to proceed quickly from theory to practice. In addition, various companies that market and make use of computational chemical methods have supported the NERMMWs, and even contributed to discussions and demonstrations therein.

## NERMMW Goals

The participant who has completed the NERMMW curriculum will have accomplished the following goals.

- Gained experience in use of PC, Macintosh, and SGI-Irix computer operating environments.
- Gained a "back-of-the-envelope", practical background in computational chemical methodology that is sufficient to allow appropriate choices of methodology for particular problems (e.g., what works for radicals, how can one predict IR spectra, *etc.*)
- Tested and gained some experience with several of the most common graphical interface driven computational chemistry software packages, in order to experience the similarities and differences in approaches among these.
- Met with industry representatives from companies marketing a number of the tested software packages, in order to make useful business contacts and to gain information useful for making eventual selections of software in course and research curricula.
- Participated in discussions aimed at discussing strategies for formulating, funding, and implementing programs in computational chemistry that are appropriate for various levels of student expertise and university resources. In particular, NERMMW holds a one-day "reprise" colloquium in the January following each June workshop, with the goal of allowing participants from all past NERMMWs to share information about successes, problems, and stratagems for utilizing computational chemistry.
- Made contact with the NERMMW region-wide group of people with common interest in making appropriate use of computational chemistry for educational purposes, in part through use of the [NERMMW WWW page](#).

## NERMMW Useful Bibliography

### *Classics*

- J. D. Roberts, "Notes on Molecular Orbital Calculations", W.A. Benjamin, New York 1962.
- Robert G. Parr, "The Quantum Theory of Molecular Electronic Structure", W. A. Benjamin, New York, NY 1964.
- Per-Olov Lowdin & Bernard Pullman, "Molecular Orbitals in Chemistry, Physics, and Biology. A Tribute to R. S. Mulliken," Academic, New York, NY 1964.
- Lionel Salem, "The Molecular Orbital Theory of Conjugated Systems", W. A. Benjamin, New York, NY 1966.
- John A. Pople & David L. Beveridge, "Approximate Molecular Orbital Theory", McGraw-Hill Book, New York, NY 1970.
- Tim Clark, "A Handbook of Computational Chemistry. A Practical Guide to Chemical Structure and Energy Calculations", Wiley, New York, NY 1985.

### *General Utility and Topical Material*

- Warren J. Hehre, Leo Radom, Paul von R. Schleyer, John A. Pople, "Ab Initio Molecular Orbital Theory", Wiley, New York, NY 1986.
- James J. P. Stewart, "MOPAC. A Semiempirical Molecular Orbital Program", *Journal of Computer-Aided Molecular Design*, **4**(1), 1-105 (1990).
- Douglas A. Smith, Editor, "Modeling the Hydrogen Bond", ACS Symposium Series 569, ACS, Washington DC 1994.
- Bruce R. Gelin, "Molecular Modeling of Polymer Structures and Properties", Hanser, New York, NY 1994.
- Paul M. Lahti, "Semiempirical Molecular Orbital Methods Applied to Polymeric Systems", <http://www.chem.umass.edu/~lahti/ARTICLE/compchem.html>.

### *Applications in the Lab and the Classroom*

- Warren J. Hehre & Janet E. Nelson, "Introducing Molecular Modeling into the Undergraduate Chemistry Curriculum", Wavefunction In., Irving, CA 1997.
- Warren J. Hehre, Alan J. Shusterman, W. Wayne Huang, "A Laboratory Book of Computational Organic Chemistry, Wavefunction Inc., Irvine CA 1996.
- James B. Foresman and ÆLeen Frisch, "Exploring Chemistry with Electronic Structure Methods", 2nd ed., Gaussian Inc., Pittsburgh, PA 1996.

## INPUT FORMATS

### Cartesian Coordinates (x,y,z)

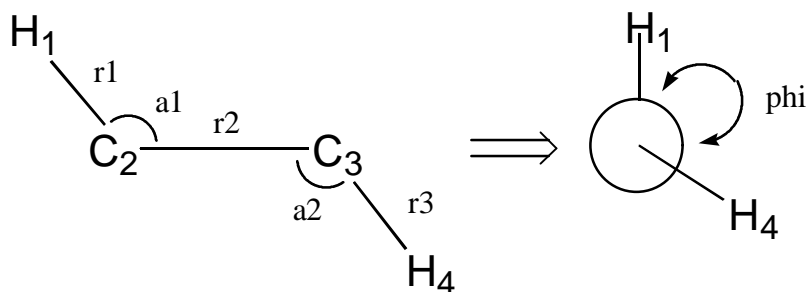
Easy to use and understand, easily exported, but many minor variations in format make these difficult to import, except for some well-established specific formats such as Protein Databank (PDB) format.

### Fractional Coordinates (a,b,c)

Very useful for crystallographic representations. Common formats such as SHELTX can be imported into some programs, but again there are difficulties with varying format

### Z-Matrix

Various formats. Very useful for partial molecular optimization and for rapid identification of bonding parameters. Not straightforward to prepare these input files, but often worth the effort for programs such as GAUSSIAN, MOPAC, AMPAC, and GAMESS.



| ATOM | DISTANCE | ANGLE | TORSION | PATH     |
|------|----------|-------|---------|----------|
| H1   |          |       |         |          |
| C2   | r1       |       |         | H1       |
| C3   | r2       | a1    |         | C2 H1    |
| H4   | r3       | a2    | phi     | C3 C2 H1 |

## MOLECULAR MECHANICS (FORCE FIELD METHODS)

### I. Basic ideas

- Use classical mechanics, not quantum mechanics
- Model a molecule by simple ball-and-spring methods
- Parameterize the model using experimental data

### II. Advantages

- Very fast methodology ( $N^2$  dependence upon molecular size)
- Good results for parameterized experimental properties
- Several well-tested and established methods
- Well-suited to modern small computers
- Best for molecular geometries, heat of formation

### III. Disadvantages

- Relies on specific parameterization of bond lengths, angles, torsions, so cannot handle new bond types without estimation
- No explicit consideration of electronic effects
- Only MMP2 and PC-Model have some consideration of conjugation electronic effects
- Does not predict any electronic properties

### IV. Some well-established methods

- MM2, MMP2 (Allinger)
- MM3 (Allinger)
- MACROMODEL (Stille)
- MMX (Gajewski, Gilbert)
- UFF (Goddard)
- CHARMM (Karplus)
- AMBER (Kollman)

# MOLECULAR MECHANICS

## Basic model

$$E_{\text{tot}} = E_{\text{steric}} + E_{\text{electronic}}$$

|                       |                             |                               |
|-----------------------|-----------------------------|-------------------------------|
| $E_{\text{steric}} =$ | $E_{\text{bond}}$           | bond stretching               |
|                       | $+E_{\text{angle}}$         | angle compression, elongation |
|                       | $+E_{\text{torsion}}$       | torsion                       |
|                       | $+E_{\text{vdw}}$           | van der Waals                 |
|                       | $+E_{\text{(corrections)}}$ | (various fudge factors)       |

## Basic components

bond stretching = Hooke's law (+ anharmonicity)

$$E = k(r-r_0)^2/2$$

$k$  = bond stretching constant

$r_0$  = "normal" bond length

angle compression = Hooke's law (+ anharmonicity)

$$E = k(\theta-\theta_0)^2/2$$

$k$  = bond stretching constant

$\theta_0$  = "normal" bond length

### torsion

$E \sim$  threefold torsional function

### van der Waals

Various formulations of "hard" and "soft" atoms

6-12 potential or modified exponentials are common

e.g.  $E = a/r^6 - b/r^{12}$

## SOME USEFUL OBSERVATIONS ABOUT MOLECULAR MECHANICS

- Methods such as MM2, MMP2, and UFF can give quantitatively excellent predictions of molecular geometries when electronic effects are not important, (e.g., alkanes, ethers, nonconjugated alkenes). This is even so for a variety of highly strained molecules, so long as modern force fields are used. Heats of formation are also predicted quite accurately for such molecules, especially with MM2 and MM3.
- For conjugated systems, MMP2 and MMX-PI can be better results than force fields that do not have explicit pi-system corrections. UFF seems to do a pretty good job for such systems.
- Inorganic systems are still pretty much *terra incognita*.
- For quite large molecules (peptides, polymers), the most effective force fields are those which model a whole group (e.g., methyl, hydroxyl) with a set of parameters. These methods include CHARMM and AMBER. Other force fields have more explicit modeling of atom-atom effects, but less effective modeling of interactions between functional groups (such as H-bonding, secondary and tertiary structural effects).
- Molecular mechanics is the most effective method with which to start a computation, even when the ultimate result is desired at a higher level of theory.
- Some useful programs using force field methods in NERMMW
  - Chem3D (Cambridgesoft, MM2, proprietary force field)
  - Hyperchem (Hypercube, various force fields available)
  - PC-Model (modified MM2/MACROMODEL)
  - Cerius<sup>2</sup> (Molecular Simulations Inc., CHARMM+, UFF)
  - Sculpt (AMBER)
  - Spartan (Merck Force Field, Sybyl)
  - Alchemy (TRIPOS force field)

# GENERAL CONCEPTS FOR COMPUTATIONAL MOLECULAR ORBITAL THEORY

I. Quantum mechanical basis is the gas phase Schrödinger equation

$$H \mathbf{y} = E \mathbf{y}$$

II. Must make practical approximations to get solutions in a reasonable time

A. Zeroth-order Methods

Simple Hückel method, Extended Hückel Method

B. Semiempirical Methods

Pariser-Parr-Pople method

Early NDO methods (CNINDO, INDO, MINDO)

Improved NDO methods (ZINDO, MNDO, AM1, PM3)

C. Ab Initio Single-Determinant Methods

Restricted Hartree-Fock

Unrestricted Hartree-Fock

Restricted Open-shell Hartree Fock

D. Multideterminant Methods

Generalized Valence Bond (GVB) method

Two-configuration SCF methods

E. Post Hartree-Fock Methods

Configuration Interaction

Perturbation methods

Higher-order coupled cluster methods

Multiconfiguration SCF methods

III. Most methods use a Linear Combination of Atomic Orbitals method

$$\Psi_{\text{LCAO}} = \sum a_i \phi_i$$

$\Psi$  = molecular orbital

$a_i$  = atomic orbital contribution to the MO

$\phi_i$  = atomic orbital basis set function

IV. Basis set selection is critical for methodological accuracy

- A. Valence basis sets -- core orbitals are parameterized as easily evaluated functions or ignored. Only the valence shell is treated explicitly.
- B. Single-zeta basis set -- all core orbitals and valence shell orbitals are explicitly treated.
- C. Double-zeta basis set -- all occupied valence shells are modeled with a double set of functions for flexibility.
- D. Split-valence basis set -- typically, only the valence shell is modeled with a double set of functions for flexibility.

V. Mathematical form of the basis set functions

- A. Almost all semiempirical basis sets are formulated using STOs (Slater-type orbitals) of the general form:

$$\Psi_{\text{STO}} = a_i e^{-r} \quad (r = \text{distance of } e^- \text{ from nucleus})$$

MO calculations involve an explicit integration across all space. STOs cannot be directly evaluated by direct methods, but since there are few of them in semiempirical calculations, they are used because they have a shape that is truer to “real” hydrogenic functions.

- B. Almost all ab initio methods use GTOs (Gaussian-type orbitals) of the general form shown below, because of their ease of integration

$$\Psi_{\text{GTO}} = a_i e^{-r^2} \quad (r = \text{distance of } e^- \text{ from nucleus})$$

However, they are not as good as STOs to represent “real” AOs.

VI. MO calculations of any sort require evaluation of a variety of integrated functions between AOs and MOs, in order to computer properties. We will give qualitative descriptions of a number of these terms for ease in discussion of MO calculations of different sorts.

- A. Overlap integrals -- evaluates the overlap between functions; more overlap equates to more interaction

$$S_{ij} = \int_0^{\infty} \phi_i \phi_j d\tau$$

- B. Resonance integrals -- evaluate interaction between orbitals, usually a bonding interaction between AOs on neighboring atoms. Proportional to the overlap integral.

$$\beta_{ij} = \beta_{ij}^0 \int_0^{\infty} \phi_i \phi_j d\tau$$

- C. Coulomb integral -- evaluates electronic attraction/repulsion between charges. Typically refers to repulsion between electrons in different orbitals.

$$J_{ij} = \int_0^{\infty} \phi_i(1)\phi_j(2) \frac{e^2}{r_{ij}} \phi_i(1)\phi_j(2) d\tau$$

- D. Exchange integral -- evaluates the effect of not being able to identify any particular electron in a real molecule. This is a completely artifactual integral that is require to rectify errors in simple LCAO methods, but its effect is profoundly important to understand bonding in real molecules with different electronic states.

$$K_{ij} = \int_0^{\infty} \phi_i(1)\phi_j(2) \frac{e^2}{r_{ij}} \phi_i(2)\phi_j(1) d\tau$$

- E. Shielding constant -- a factor representing the amount of shielding of the atomic nucleus by (n-1) electrons, relative to the n<sup>th</sup> electron. Such shielding constants are built into computations in a variety of ways, but are seldom explicitly seen or manipulated by the user.

VII. Typical MO calculations carry out an iterative fit to self-consistency by the following procedure:

- Formulate a guess as to the final distribution of electrons and orbitals in the calculation.
- Apply the guess to a specific molecular hamiltonian, and solve for the appropriate eigenvectors/eigenvalues from the guess.
- Use the result as the next guess, and repeat until self-consistent.

VIII. The various molecular orbital methods are most clearly differentiated by their formulation of a molecular hamiltonian expression. Typical terms that may be treated in the hamiltonian include:

- Nuclear-nuclear repulsion  
(typically uses a coulombic type interaction)
- Electron-nuclear attraction  
(typically uses a coulombic-type interaction)
- Electron-electron repulsion  
(various schemes, very difficult to estimate, impossible to solve)
- Spin-orbit coupling between electrons, nuclei (typically ignored)
- Relativistic effects  
(typically ignored, except for electron spin correlation effects)
- Time dependent effects  
(typically ignored, except for trajectory and spectroscopic studies)

IX. Computational methodology does affect how a computation is actually done. Most calculations are done using Fock's modification of the so-called Hartree operator, in place of the more complex Schrödinger hamiltonian operator. The basic idea is that the Fock hamiltonian is written as a sum of effective one-electron terms. This is tantamount to allowing interaction between atomic orbitals in an LCAO type of fashion, rather than trying to solve the insoluble multi-body problem of the strict hamiltonian.

$$F(1,2,\dots,n) = \sum_i \left[ -\frac{1}{2} \nabla_i^2 + V(p) \right]$$

where  $V(p)$  is an one-electron potential energy function of some sort. The Fock operator  $F$  replaces the hamiltonian operator  $H$ .

X. The Fock operator basically treats the problem of the  $n^{\text{th}}$ -electron circling a nucleus that is shielded by  $(n-1)$  electrons, and can be set up to solve iteratively for all  $n$  electrons in a self-consistent sense. However, a final problem that must be solved is that electrons are indistinguishable, and cannot be labeled. The actual molecular wavefunction must reflect this property. In addition, relativistic Dirac treatment of molecular wavefunctions requires the spin quantum number. Finally, the Pauli principle requires that overall molecular wavefunctions be overall antisymmetric to electron label interchange. The sum of all these principles makes the actual molecular wavefunction far more imposing than is usually considered by the nonexpert.

$$\Psi_{\text{overall}} = \Psi_{\text{space}} \Psi_{\text{spin}}$$

$$\Psi_{\text{overall}}(1,2,3,\dots,n) = -\Psi_{\text{space}}(2,1,3,\dots,n)$$

$$\Psi_{\text{overall}} = \frac{1}{N}(\phi_i + \phi_j) \bullet (\alpha(1)\beta(2) - \alpha(2)\beta(1))$$

but this gets messy really fast for more than two electrons, so one uses matrix math to overcome the problem by expressing the wavefunction as a Slater determinant.

$$\Psi(1,2,3,\dots,n) = N \bullet \begin{vmatrix} \psi_1(1)\alpha(1) & \psi_1(1)\beta(1) & \psi_2(1)\alpha(1) & \psi_n(1)\beta(1) \\ \psi_1(2)\alpha(2) & \psi_1(2)\beta(2) & \psi_2(2)\alpha(2) & \dots \\ \psi_1(3)\alpha(3) & \dots & \dots & \dots \\ \psi_1(2n)\alpha(2n) & \dots & \dots & \psi_n(1)\beta(1) \end{vmatrix}$$

One may use standard matrix mathematics to manipulate such wavefunctions, which fulfill the desiderata described above.

## XI. Basic things to look for in a molecular orbital calculation

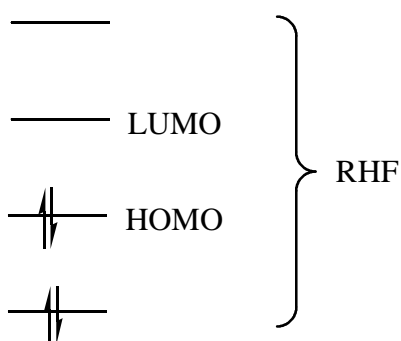
- SCF convergence ( $\Delta E_{\text{SCF}}$  approaches zero as a limit)
- Density convergence (atomic charges and AO populations stop changing)
- Gradient convergence (molecule is in a true energy minimum, with all positive 2nd derivatives = vibrational frequencies)

## XII. What do you want to get from the calculation?

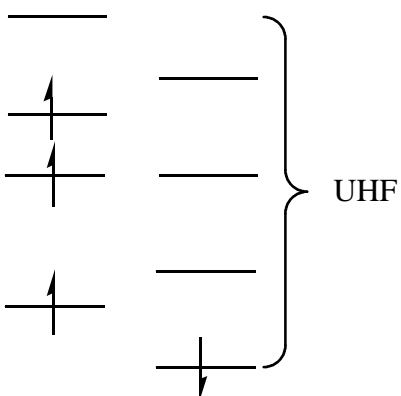
- Molecular geometry
- Overall molecular energy (heat of formation)
- Molecular orbital energies (eigenvalues = orbital ionization potentials)
- Molecular orbital appearances
- Charge and/or spin distributions
- Dipole moments
- Vibration frequencies

XII. Semiempirical calculations scale as  $\sim N^3$  with molecular size  
Ab initio SCF calculations scale as  $\sim N^4$  with molecular size  
Post Hartree-Fock calculations scale as  $\sim N^{5-6}$  molecular size

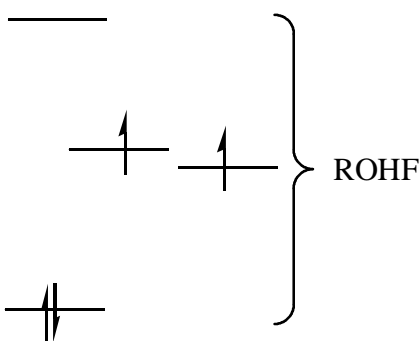
## CHOOSING THE RIGHT WAVEFUNCTION



Most molecules are closed-shell systems with all electrons paired in molecular orbitals by the Aufbau principle. For these systems, restricted Hartree-Fock wavefunctions are appropriate (RHF). This is the simplest and fastest type of SCF computation.



Open-shell molecules are most often treated by the unrestricted Hartree-Fock method (UHF) developed by Pople and Nesbet.  $\alpha$ - and  $\beta$ -electrons are treated separately. This method gives good results for geometry, and has some basic inclusion of electron correlation effects. However, this wavefunction does not give proper treatment of the spin expectation value for a molecule [ $\langle S^2 \rangle = S(S+1)$ ], a problem called spin contamination.



Sometimes open-shell systems are treated by the Roothaan restricted open-shell Hartree-Fock method (ROHF). This method is rather more complex than UHF, since the interactions between the open shell and the closed shell must be properly treated. Correlation effects are not well treated. But, there is no spin contamination problem. Sometimes, an ROHF energy is obtained using a UHF geometry.

## PRACTICAL CONSIDERATIONS FOR SEMIEMPIRICAL CALCULATIONS: TRUST BUT VERIFY

- Build and pre-optimize your structure with a force-field method. Consider carefully whether you should use symmetry, or intentionally break symmetry because you expect a non-symmetrical structure. Alternatively, use a known molecular structure from X-ray data or other information (e.g., force field geometry).
- Choose a proper wavefunction and hamiltonian, and optimize geometry at the semi-empirical level. Follow the standard methods known for that level of theory, unless you have already found that they do not work for your problem.
- Wherever possible, carry out calculations for one or more test molecule whose properties of interest are known experimentally. Use the same procedure you intend to use for your unknown molecules of interest. These are your procedural benchmarks.
- Bigger is not always better. A variety of molecular properties are not only more readily determined by semiempirical than by ab initio methods, but are also more accurately determined than is accomplished by small basis set ab initio methods.
- Always be ready to mistrust somewhat any single semiempirical result. Such methods are not governed by the variational principle, so one cannot count on doing more complex semiempirical methods to give improved answers. Always verify by comparison to experiment or to high level ab initio results wherever possible.

## SEMIEMPIRICAL METHODS

### I. Zeroth-order methods --

#### A. $\pi$ -Hückel approach

- 2p- $\pi$  only basis set
- $\beta_{ij}$  controlled only by connectivity
- $S_{ij}=0$  for all i,j
- No consideration of geometric, steric effects
- No consideration of coulombic or exchange effects

#### Utility

- General nodal properties of  $\pi$ -MOs
- General symmetry properties of  $\pi$ -MOs
- Qualitative evaluation of relative HOMO-LUMO splitting
- Qualitative evaluation of energy differences between states with different orbital occupancies

#### B. Extended-Hückel approach

- Valence s,p and s,p,d basis sets
- $\beta_{ij}$  parameterized by diatomic overlap and distance
- $S_{ij}$  explicitly considered
- No direct consideration of steric energy effects
- No consideration of coulombic or exchange effects

#### Utility

- General symmetry, nodal properties of frontier-MOs
- Semiquantitative effects of geometry/overlap on MO energies and nodal characteristics
- Semiquantitative evaluation of relative HOMO-LUMO splitting
- Semiquantitative evaluation of energy differences between states with different orbital occupancies

## II. Pariser-Parr-Pople (PPP) method

- $2p-\pi$  only basis set
- $\beta_{ij}$  parameterized by diatomic overlap and distance to yield proper geometries or orbital energies
- $S_{ij}$  explicitly considered for  $p-\pi$  orbitals only
- Coulomb integrals parameterized as a function of diatomic overlap and distance using either an approximate coulomb term or an artificial distance function (Mataga integrals), parameterized to yield proper geometries or orbital energies ( $\gamma_{ij}$ )
- No consideration of exchange effects

Good results for UV-vis spectral predictions of systems with no  $n \rightarrow \pi^*$  transitions, for the Hinze-Beveridge spectral parameterization.

## II. Complete Neglect of Differential Overlap (CNDO) method

- Valence STO basis set
- $\beta_{ij}$  parameterized in manners similar to those of related CNDO/1, CNDO/2, CNDO/S methods
- $S_{ij}$  explicitly considered
- Coulomb integrals parameterized as a function of diatomic overlap and distance using either an approximate coulomb term or an artificial distance function (CNDO/S Mataga integrals), parameterized to yield proper geometries or orbital energies ( $\gamma_{ij}$ )
- AO core energies determined using experimentally weighted ionization potentials.
- No consideration of exchange effects.

Pretty good results for geometry optimizations. Ionization potentials usually too large. Cannot distinguish between states of different multiplicity with the same orbital occupancy. Good results for UV-vis spectral predictions for the Jaffé CNDO/S parameterization.

## III. Intermediate Neglect of Differential Overlap (INDO) method

- Valence STO basis set
- $\beta_{ij}$  parameterized by diatomic overlap and distance to yield proper geometries (CNDO/1, CNDO/2) or orbital energies (CNDO/S)
- $S_{ij}$  explicitly considered

- Coulomb integrals parameterized as a function of diatomic overlap and distance using modified coulomb term or an artificial distance function (INDO/S Mataga, Ohno-Klopman, Warshol integrals), parameterized to yield proper geometries or orbital energies ( $\gamma_{ij}$ )
- AO core energies determined using experimentally weighted ionization potentials.
- Monatomic exchange only is considered (exchange between two electrons on the same atom).

Pretty good results for geometry optimizations. Ionization potentials usually too large. Pretty good results for relative energies of states with the same orbital occupancy but different multiplicity. Very good prediction of spin densities for simple organics. Very good results for UV-vis spectral predictions using the Zerner ZINDO plus configuration interaction parameterization and procedure

#### IV. Modified Intermediate Neglect of Differential Overlap (MINDO)

- See J. J. P. Stewart, *J. Comp. Mol. Design*, **4**(1), 1-105 (1990) for a brief summary of methodology in MOPAC with references.

Work of Michael J. S. Dewar. A first effort to modify Pople's INDO scheme to give all properties of molecules with one set of parameters. Major advance was the separate parameterization of coulomb integrals as s-s, s-p, and p-p types, giving more parametric flexibility. MINDO/3 can still be found in the non-graphical program MOPAC, but has been largely supplanted. MINDO/3 is remarkably good for organic carbocation computations, but has a variety of other deficiencies that led to its being replaced by more sophisticated treatments. Its major drawback is that parameters (e.g., resonance integrals) were formulated as diatomics, not monatomics. Thus, one has to have parameters for a large variety of diatomic combinations, rather than a set of atoms.

## V. Modified Neglect of Diatomic Differential Overlap (MNDO)

- See J. J. P. Stewart, *J. Comp. Mol. Design*, **4(1)**, 1-105 (1990) for a brief summary of methodology in MOPAC with references.

Work of Michael J. S. Dewar and James J. P. Stewart. Used many of the strategies of MINDO, but redone with monatomic parameters to increase flexibility. The first semi-empirical method to gain wide usage for a variety of organic structural problem analyses, due to its incorporation of geometry optimization methods. Good for geometries, ionization potentials. OK for dipole moments, spin densities. Problems with small rings and with neighboring lone-pair to lone-pair interactions (e.g., hydrazine, hydrogen peroxide). Problems with transition state analysis.

## VI. Austin Model 1 (AM1)

- See J. J. P. Stewart, *J. Comp. Mol. Design*, **4(1)**, 1-105 (1990) for a brief summary of methodology in MOPAC with references.

Work of Michael J. S. Dewar. Basic reparameterization of MNDO with different statistical breakdown of parameters. The de facto standard for semiempirical computations of up to 500 atoms. Very good for geometries, ionization potentials. Good for dipole moments. OK for spin densities. Some problems with small rings and with neighboring lone-pair to lone-pair interactions (e.g., hydrazine, hydrogen peroxide). Transition state analysis better than MNDO.

## VI. MNDO-PM3

- See J. J. P. Stewart, *J. Comp. Mol. Design*, **4(1)**, 1-105 (1990) for a brief summary of methodology with references.

Work of J. J. P. Stewart. Radically different method for reparameterization of MNDO a large number of experimental molecules and data. Less tested than AM1. Very good for geometries, ionization potentials. Good for dipole moments. OK for spin densities. Some problems with small rings and with neighboring lone-pair to lone-pair interactions (e.g., hydrazine, hydrogen peroxide). Transition state analysis better than MNDO. Particularly recommended for hypervalent species (e.g., pentavalent phosphorus).

## CONFIGURATION INTERACTION

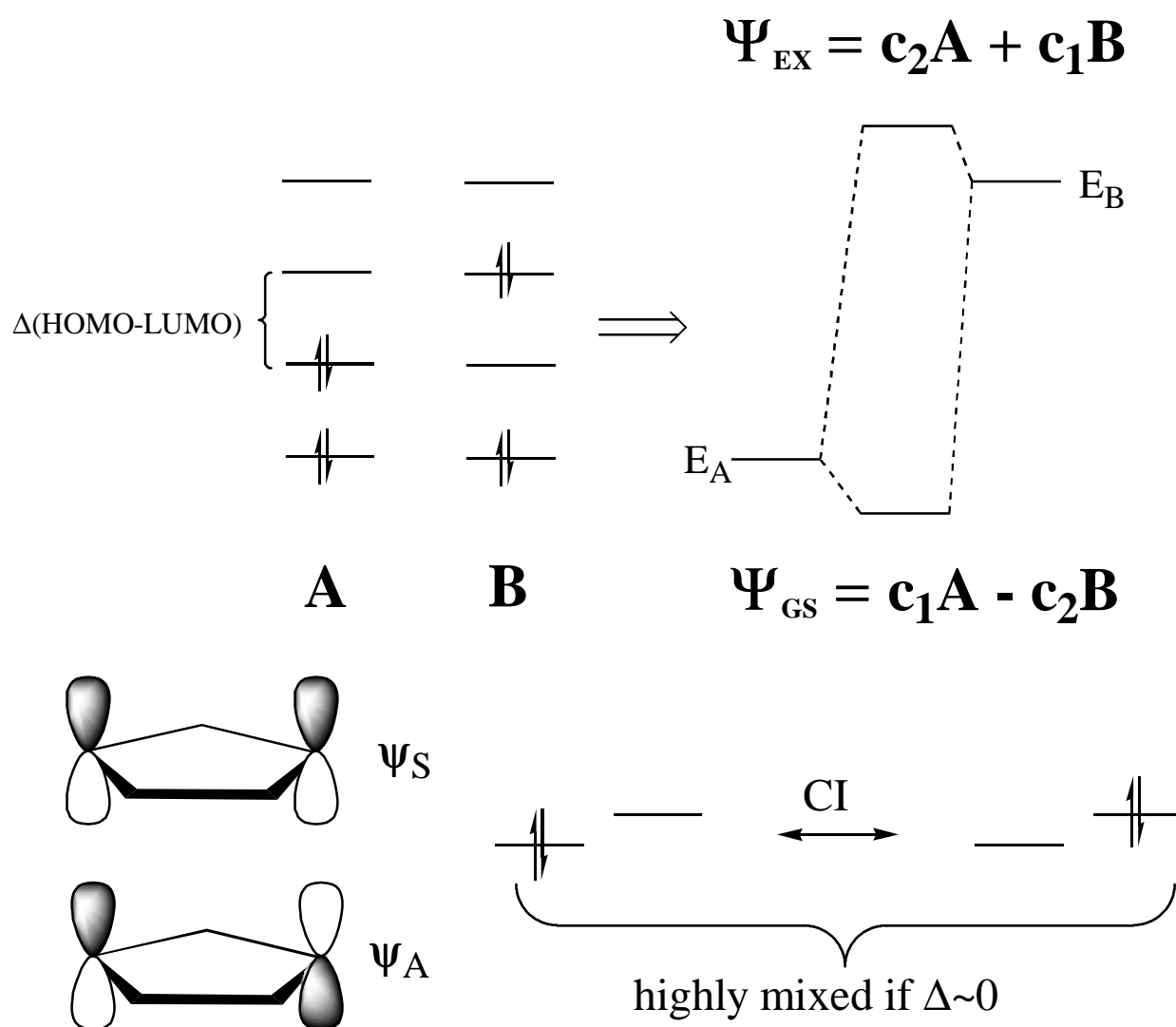
CI is used to mix excited state with ground state configurations to get a better representation of the “true” molecular wavefunction. The variational theorem states that lower energy wavefunctions are always better wavefunctions for an ab initio hamiltonian. Since MO theory takes poor account of electron correlation effects, a MO-CI combination method makes up for this deficiency. Unfortunately, MO-CI calculations can take a lot of time (since many configurations can be generated for large molecules) and can be difficult to understand, since one loses the (simplistic) concept of one configuration corresponding to one state. The overall geometries, energies, and molecular properties are improved, however, relative to those of simple SCF-MO calculations. As a bonus, excited states are far better described by MO-CI methods than by SCF-MO methods.

CI calculations are often done by generating a limited number of configurations (limited CI) from a limited number of molecular orbitals (active space MOs). The dangers are: (i) what is the best set of orbitals to use? (ii) how many configurations is enough for a good description? This is referred to as the “size inconsistency problem”.

| GROUND STATE | SINGLE EXCITATIONS |    |    | DOUBLE EXCITATIONS |    |
|--------------|--------------------|----|----|--------------------|----|
| —            | —                  | —  | ↑  | —                  | ↑  |
| —            | ↑                  | —  | —  | ↑↓                 | ↓  |
| ↑            | —                  | ↑↓ | ↑  | ↑                  | ↑  |
| ↑↓           | ↑↓                 | ↑  | ↓  | —                  | —  |
| ↑↓           | ↑↓                 | ↑↓ | ↑↓ | ↑↓                 | ↑↓ |

## CONFIGURATION INTERACTION RESULTS

CI lowers the energy of a molecular state relative to its main configuration. Configurations are not real, states are real. Sometimes a state may be a highly mixed array of configurations, that a single configuration description is impossible: this is often so for excited states. Such situations occur when the typical MO description of a molecule is poor. Diradical closed-shell singlet states are a classic example of such a situation, in which the state cannot be described by a single Lewis structure ( $\Delta \sim 0$ ,  $c_1 \sim c_2 \sim 0.707 = \frac{1}{\sqrt{2}}$ .)



## POST HARTREE-FOCK PERTURBATION THEORY

For any state that is fairly closely described by a single configuration (single-determinant) wavefunction, perturbation theory may be applied. The brief summary below is the Møller-Plesset (MP) perturbation treatment, and is not unique, but it is one of the more popular perturbation formulations due to its use in the *Gaussian* suite of programs.

Basically, the MP method adds corrections to a zeroth-order hamiltonian, which in this case is the one-electron Fock operator used for SCF wavefunctions. One can express an exact, full CI wavefunction as a power series expansion that may be truncated as desired, viz.

$$\Psi_{\text{MP}} = \Psi(0) + \lambda_1\Psi(1) + \lambda_2\Psi(2) + \lambda_3\Psi(3) + \dots$$

MP2 is truncated after the second-order term, MP3 after the third order term, and so on. After truncation, the wavefunction is subjected to SCF optimization of the appropriate coefficients. Typically, MP4 is the highest level used, although fifth-order treatments have been described. The size of an MP4 calculation can scale as  $N^{5-6}$ , so this is not undertaken lightly.

A great advantage of the perturbation methods is that they are “size consistent”. Unlike the case for limited CI calculations, in MP calculations one does not suffer from inconsistent truncation imprecision for large molecules by comparison to small molecules. This makes them very popular for comparing properties of structurally quite different molecules.

A disadvantage of the MP methods is that they do not treat excited states well, and do a very poor job for wavefunctions that are not well-approximated by a single configuration. Interestingly, although perturbation wavefunctions have been observed to be “well-behaved” in almost all cases, they are not variational.