Semiemprical Methods

Chapter 5

Semiemprical Methods

- Because of the difficulties in applying ab initio methods to medium and large molecules, many semiempirical methods have been developed to treat such molecules.
- The earliest semiempirical methods treated only the π electrons of conjugated molecules.
- In the **π-electron approximation**, the n_π π electrons are treated separately by incorporating the effects of the σ electrons and the nuclei into some sort of effective π-electron Hamiltonian $\hat{H}_\pi$ (recall the similar valence-electron approximation

$$\hat{H}_\pi = \sum_{i=1}^{n_\pi} \hat{H}_\pi^{\text{core}}(i) + \sum_{i=1}^{n_\pi} \sum_{j>i} \frac{1}{r_{ij}}$$

$$\hat{H}_\pi^{\text{core}}(i) = -\tfrac{1}{2}\nabla_i^2 + V(i)$$

- where V_i is the potential energy of the ith π electron in the field produced by the nuclei and the σ electrons. The *core* is everything except the π electrons.

- The most celebrated semiempirical π-electron theory is the Hückel molecular-orbital (HMO) method, developed in the 1930s to treat planar conjugated hydrocarbons.
- Here the π-electron Hamiltonian is approximated by the simpler form

$$\hat{H}_{\pi} = \sum_{i=1}^{n_{\pi}} \hat{H}^{\text{eff}}(i)$$

- where $\hat{H}^{\text{eff}}(i)$ incorporates the effects of the π-electron repulsions in an average way.
- In fact the Hückel method does not specify any explicit form for $\hat{H}^{\text{eff}}(i)$.
- Since the Hückel π-electron Hamiltonian is the sum of one-electron Hamiltonians, a separation of variables is possible.

We have

$$\hat{H}_\pi \psi_\pi = E_\pi \psi_\pi$$

$$\psi_\pi = \prod_{i=1}^{n_\pi} \phi_i$$

$$\hat{H}^{\text{eff}}(i)\phi_i = e_i \phi_i$$

$$E_\pi = \sum_{i=1}^{n_\pi} e_i$$

- The wave function takes no account of spin or the antisymmetry requirement.
- To do so, we must put each electron in a spin-orbital $u_i=\phi_i\sigma_i$. The wave function ψ_π is then written as a Slater determinant of spin-orbitals.
- Since $\hat{H}^{\text{eff}}(i)$ is not specified, there is no point in trying to solve above eq. directly.
- Instead, the variational method is used.

The next assumption in the HMO method is to approximate the π MOs as LCAOs.

$$\phi_i = \sum_{r=1}^{n_C} c_{ri} f_r$$

- This is a linear variation function, the optimum values of the coefficients for the c_{ri} lowest π MOs satisfy

$$\sum_{s=1}^{n_C} [(H_{rs}^{\text{eff}} - S_{rs} e_i) c_{si}] = 0, \qquad r = 1, 2, \ldots, n_C$$

- where the e_i's are the roots of the secular equation

$$\det(H_{rs}^{\text{eff}} - S_{rs} e_i) = 0$$

The key assumptions in the Hückel theory involve the integrals in the above equation.

$$H_{rr}^{\text{eff}} = \int f_r^*(i)\hat{H}^{\text{eff}}(i)f_r(i)\,dv_i \equiv \alpha$$

$$H_{rs}^{\text{eff}} = \int f_r^*(i)\hat{H}^{\text{eff}}(i)f_s(i)\,dv_i \equiv \beta \quad \text{for } C_r \text{ and } C_s \text{ bonded}$$

$$H_{rs}^{\text{eff}} = 0 \quad \text{for } C_r \text{ and } C_s \text{ not bonded together}$$

$$S_{rs} = \int f_r^*(i)f_s(i)\,dv_i = \delta_{rs} \qquad f_r \equiv C_r 2p\pi$$

- where δ_{rs} is the Kronecker delta. a is called the **Coulomb integral** and β is the **bond integral** (or resonance integral).
- The HMO secular determinant equals the number of conjugated atoms. Students sometimes make the error of assuming that this order always equals the number of π electrons.

- Hückel theory neglects interelectronic repulsions,
- A semiempirical π-electron theory that takes electron repulsion into account and thereby improves on the Hückel method is the **Pariser–Parr–Pople (PPP) method**, developed in 1953.

$$\hat{H}_\pi = \sum_{i=1}^{n_\pi} \hat{H}_\pi^{\text{core}}(i) + \sum_{i=1}^{n_\pi} \sum_{j>i} \frac{1}{r_{ij}}$$

$$\hat{H}_\pi^{\text{core}}(i) = -\tfrac{1}{2}\nabla_i^2 + V(i)$$

- The π-electron Hamiltonian $\hat{H}_\pi$ has the same form as the all-electron operator

The Roothaan equations become

$$\sum_s c_{si}(F_{\pi,rs} - \varepsilon_i S_{rs}) = 0, \qquad r = 1, \ldots, b$$

$$F_{\pi,rs} = H_{\pi,rs}^{\text{core}} + \sum_{t=1}^{b} \sum_{u=1}^{b} P_{tu}\left[(rs|tu) - \tfrac{1}{2}(ru|ts)\right]$$

- In addition to assuming σ-π separability, the PPP method makes further approximations. As in Hückel theory, overlap is neglected:

$$S_{rs} \equiv \langle f_r(1) | f_s(1) \rangle = \delta_{rs}$$

- Consistent with the neglect of overlap integrals, when evaluating electron-repulsion integrals the PPP method makes the approximation of **zero differential overlap** (ZDO):

$$[f_r(1)]^* f_s(1)\, dv_1 = 0, \qquad \text{for } r \neq s$$

Thus

$$[f_r(1)]^* f_s(1)\, dv_1 = 0, \qquad \text{for } r \neq s$$

$$(rs|tu) = \langle f_r(1) f_t(2) | 1/r_{12} | f_s(1) f_u(2) \rangle$$

$$(rs|tu) = \delta_{rs}\delta_{tu}(rr|tt) = \delta_{rs}\delta_{tu}\gamma_{rt}$$

electron-repulsion integrals

- Thus the method ignores many (but not all) of the electron-repulsion integrals, thereby greatly simplifying the calculation.
- In particular, all three- and fou-rcenter electron-repulsion integrals are ignored. The ZDO approximation is not used in the $H_{\pi,rs}^{\text{core}}$ integrals.

General Semiempirical Methods

- The HMO and PPP methods apply only to planar conjugated molecules and treat only the π electrons.
- The semiempirical MO methods discussed in this section apply to all molecules and treat all the valence electrons.
- Semiempirical MO theories fall into two categories: those using a Hamiltonian that is the sum of one-electron terms, and those using a Hamiltonian that includes two-electron repulsion terms, as well as one-electron terms.
- The Hückel method is a one-electron theory, whereas the Pariser–Parr–Pople method is a two-electron theory.

The Extended Huckel Method

- The most important one-electron semiempirical MO method for nonplanar molecules is the **extended Hückel theory**.
- The extended Hückel (EH) method begins with the approximation of treating the valence electrons separately from the rest.
- The valence-electron Hamiltonian is taken as the sum of one-electron Hamiltonians:

$$\hat{H}_{\text{val}} = \sum_i \hat{H}_{\text{eff}}(i)$$

$$\phi_i = \sum_{r=1}^{b} c_{ri} f_r$$

$$\hat{H}_{\text{eff}}(i)\phi_i = e_i\phi_i$$

$$E_{\text{val}} = \sum_i e_i$$

$$\det(H_{rs}^{\text{eff}} - e_i S_{rs}) = 0$$

$$\sum_s [(H_{rs}^{\text{eff}} - e_i S_{rs})c_{si}] = 0, \qquad r = 1, 2, \ldots, b$$

- All this is similar to simple Hückel theory. However, the extended Hückel theory does not neglect overlap. And other integrals ….

The CNDO, I NDO, and NDDO Methods

- Several semiempirical two-electron MO generalizations of the PPP method were developed that apply to both planar and nonplanar molecules.
- The **complete neglect of differential overlap (CNDO) method** was proposed by Pople, Santry, and Segal in 1965.
- The **intermediate neglect of differential overlap (INDO) method** was proposed by Pople, Beveridge, and Dobosh in 1967.
- The **neglect of diatomic differential overlap** (NDDO) method suggested by Pople, Santry, and Segal in 1965, is an improvement on INDO in which differential overlap is neglected only between AOs centered on different atoms:
- $f^*_r(1)\, f_s(1)\, dv_1 = 0$ only when AOs *r* and *s* are on different atoms.
- these methods treat only the valence electrons explicitly.
- The valence-electron Hamiltonian is:

$$\hat{H}_{\text{val}} = \sum_{i=1}^{n_{\text{val}}} \left[-\tfrac{1}{2}\nabla_i^2 + V(i) \right] + \sum_{i=1}^{n_{\text{val}}} \sum_{j>i} \frac{1}{r_{ij}} \equiv \sum_{i=1}^{n_{\text{val}}} \hat{H}_{\text{val}}^{\text{core}}(i) + \sum_{i=1}^{n_{\text{val}}} \sum_{j>i} \frac{1}{r_{ij}}$$

$$\hat{H}_{\text{val}}^{\text{core}}(i) \equiv -\tfrac{1}{2}\nabla_i^2 + V(i)$$

The MNDO, AM1, PM3, PM6, PM6-D3H4, PM7, and R M1 Methods

- Pople's aim in the CNDO and INDO methods was to reproduce as well as possible the results of minimal-basis-set ab initio SCF MO calculations with theories requiring much less computer time than ab initio calculations.
- Since CNDO and INDO use approximations, we can expect their results to be similar to but less accurate than minimal-basis ab initio SCF MO results.
- Thus these methods do pretty well on molecular geometry but fail for binding energies.
- Dewar and co-workers devised several semiempirical SCF MO theories that closely resemble the INDO and NDDO methods.
- However, Dewar's aim was not to reproduce ab initio SCF wave functions and properties but to have a theory that would give molecular binding energies with chemical accuracy within 1 kcal/mol and that could be used for large molecules without a prohibitive amount of calculation.
- It might seem unlikely that one could devise an SCF MO theory that involves approximations to the ab initio Hartree–Fock method but that succeeds for binding energies, where the Hartree–Fock theory fails.
- However, by proper choice of the parameters in the semiempirical SCF theory, one can actually get better results than ab initio SCF calculations, because the choice of suitable parameters can compensate for the partial neglect of electron correlation in ab initio SCF theory.

- The semiempirical theories of this section, which follow Dewar's approach to parametrization, will be called *Dewar-type theories*.
- These theories treat only the valence electrons, and most of these theories use a minimal-basis set of valence Slater-type *s* and *p* AOs (with orbital exponents given values determined by parametrization) to expand the valence-electron MOs.
- The Fock–Roothaan equations (with the overlap integrals S_{rs} taken as δ_{rs}) are solved to find semiempirical SCF MOs.
- Some degree of neglect of differential overlap is used to eliminate many of the electron-repulsion integrals.
- In ab initio methods, the integrals occurring in the Fock matrix elements F_{rs} are evaluated accurately, but this is not the approach used in Dewar-type theories.
- Dewar-type theories take the one-center electron-repulsion integrals (ERIs) as parameters whose values are chosen to fit experimental atomic energy-level data and calculate the two-center ERIs from the values of the one-center ERIs and the internuclear distances using an approximate formula that may involve parameters.
- The remaining integrals are evaluated from approximate parameter containing formulas that are designed not to give values that accurately reproduce ab initio values but to be consistent with the approximations used in the theory.
- The Dewar-type theories are parametrized so as to yield good values of the 25 °C *gas phase* standard enthalpy of formation $\Delta H^\circ_{f,298}$.

The first useful Dewar-type theory was the MINDO/3 (third version of the modified INDO) method, published in 1975.

- MINDO/3 is based on the INDO approximation.
- Because MINDO/3 did not meet Dewar's aims, Dewar and Thiel developed the **MNDO** (modified neglect of diatomic overlap) method.
- MNDO gives substantially improved results as compared with MINDO/3.
- In 1985 Dewar and co-workers published an improved version of MNDO called **AM1** (Austin model 1, named for the University of Texas at Austin).
- AM1 has been parametrized for nearly all the main-group elements and for Zn, Cd, and Hg.
- In 1989, Stewart re-parametrized AM1 to give the **PM3 method** (parametric method 3, methods 1 and 2 being MNDO and AM1)
- The **RM1 method** (Recife Model 1, so named because it was developed at the Federal University of Pernambuco in Recife, Brazil) has exactly the same structure as AM1, but all 191 parameters for the atoms C, H, O, N, S, P, F, Cl, Br, I were reevaluated using data from 1736 molecules (as compared with about 200 molecules used for AM1)

Since parameters are available for only 10 elements, RM1 is less widely applicable than AM1 or PM3.

- The **PDDG/PM3** and **PDDG/MNDO** methods are modifications of PM3 and MNDO
- add a certain function, called the pairwise distance directed Gaussian (PDDG) function, containing additional parameters, to the core-repulsion function, thereby significantly increasing the accuracy of these methods
- A major limitation of the original versions of the MNDO, AM1, and PM3 methods is that they use a basis set of *s* and *p* valence AOs only, so they cannot be used with transition metal compounds. (In Zn, Cd, and Hg, the *d* electrons are not valence electrons.)
- Moreover, for compounds containing such second-row elements as S, the contributions of *d* orbitals to MOs are significant, and these methods do not perform well for such compounds.
- Thiel and Voityuk extended MNDO to include *d* orbitals for many second-row and later elements, giving the **MNDO/d** method. MNDO/d does not add *d* orbitals for first-row elements, so for a compound containing only C, H, O, and N, MNDO/d is precisely the same as MNDO.
- Stewart revised the PM3 method to give the **PM5 method** (Stewart, MOPAC2002), which gives more accurate $\Delta H°_{f,298}$ values than PM3 and has been parametrized for 50 elements, including many transition elements. (The PM4 method was never published, nor was it made available in a program. PM5 was never published.)

In 2007, Stewart published the **PM6 method**, which has been parametrized for 70 elements (nearly all the main-group elements and nearly all the transition elements).

- Data from 9000 compounds were used in the parametrization, including both experimental data and data from HF/6-31G* and B3LYP/6-31G* calculations.
- PM6 gives a significant improvement in accuracy over its predecessors.
- Semiempirical methods usually give satisfactory bond lengths and bond angles, but their results are not as accurate as ab initio or DFT results with a suitable-size basis set.
- The semiempirical theories discussed so far do poorly in dealing with noncovalent interactions (hydrogen bonding, dispersion) that are key to determining the structures of biological molecules. The addition of parameter-containing empirical terms to the PM6 energy that correct for dispersion (D) and hydrogen bonding (H) gave the **PM6-DH2** method.
- Further improvements in the dispersion and hydrogen-bonding correction terms gave the **PM6-DH+** and **PM6-D3H4** methods.
- In 2012, Stewart published the **PM7 method**, a reparametrization of PM6 that improves the accuracy of heats of formation of organic compounds and that includes dispersion and hydrogen-bonding terms to better represent intermolecular interactions.

Semiempirical methods are widely available in many programs.

- *Gaussian* 09: includes the MNDO, AM1, PM3, PPDG, PM6, MINDO/3, INDO, and
- CNDO methods.
- Spartan: includes the MNDO, MNDO/d, AM1, PM3, PM6, and RM1 methods.
- HyperChem: has the MNDO, MNDO/d, AM1, PM3, RM1, MINDO/3, CNDO, INDO, INDO/S and extended Hückel methods.
- MOPAC2012: has the MNDO, AM1, PM3, PM6, PM6-DH2, PM6-DH+, PM7, and RM1 methods.
- An older version is MOPAC2009
- AMPAC 9: has the SAM1, AM1, PM3, RM1, PM6, MNDO, and MNDO/d methods