ence in the enzyme from different species or tissues could be detected with this type as inhibitor. Adjacent to the active site, but not part of the active site, is a hydrophobic bonding region. This latter region has undergone considerable evolutionary change between vertebrate sources, bacteria, and the T_2-phage coded enzyme. Therefore, inhibitors that partly complex in the active site and partly complex in the hydrophobic bonding region can be constructed that show huge differences in binding between mammalian enzymes on the one hand and bacterial or protozoal enzymes on the other; these binding differences are sufficiently large for chemotherapeutic utility. Small differences in the hydrophobic bonding region between the enzyme from L1210 mouse leukemia and normal mouse liver can be detected, but these are too small for therapeutic use. These small differences can be greatly magnified with active-site-directed irreversible inhibitors composed of three moieties: a part that complexes in the active site, a part that complexes with the hydrophobic bonding region, and a part that can form a covalent bond in a hydrophilic region of the enzyme not in the active site. The formation of this covalent bond leads to rapid inactivation of the enzyme. Inhibitors can be constructed that can destroy the L1210 mouse leukemia enzyme but not destroy the enzyme from normal mouse liver, spleen, or intestine—so-called isozyme specificity. Since large differences in the ability of these inhibitors to diffuse through the L1210 cell wall have been observed, current research revolves around modifying those structures with good transport to give isozyme specificity.

All-Valence-Electron Semiempirical Self-Consistent Field Calculations

H. H. JAFFÉ

Department of Chemistry, University of Cincinnati, Cincinnati, Ohio 45221

Since the early 1930's, chemists have been interested in the quantum mechanical methods opened up by the work of Schrödinger, Heisenberg, and Dirac, particularly in their applications to the calculation of the electronic structure of molecules. Of the several approaches to molecular calculation that have evolved, the molecular orbital method has proved the most flexible, and has virtually displaced all others. From the early days, molecular orbital calculations have been approached in two independent ways. Both methods have almost universally been based on two assumptions: that the many-electron wave functions of the molecule required by the theory can be factored into independent one-electron functions, molecular orbitals, each of which depends only on the coordinates of one electron; and that each of these molecular orbitals (MO's) can be expressed as a linear combination of atomic orbitals, the LCAO approximation.

The first of the two approaches is an *a priori* (or *ab initio*) calculation, *i.e.*, calculations on molecules using only a limited number of fundamental constants as external input to the calculations, the charge and mass of the electron, e and m, Planck's constant, h, and the velocity of light, c. Such *a priori* calculations look relatively straightforward but always run into problems of extreme computational difficulties, even for relatively small molecules.

Under the pressure of attempting to obtain chemically interesting information short of a complete *a priori* calculation, work started very early on *semiempirical* calculations. Although even *a priori* calculations involve many approximations, the semiempirical methods require many further, sometimes extremely crude, and often unjustifiable approximations. The effect of these approximations is minimized by choosing numerical values for some or all integrals arising in the calculations by fitting results to known experimental data, and then using these same values for other calculations.

One of the earliest triumphs of this approach was the recognition that, to a fair approximation, it is possible to separate the σ and π electron systems in planar unsaturated and aromatic organic molecules, *i.e.*, to treat the electron system which has a nodal plane coincident with the molecular plane separately from the other electrons. This separation has led to the well-known Hückel molecular orbital (HMO) method.[1] The success of this method is so well known that no further discussion is needed here.

The HMO method involves, among many others, one particularly onerous approximation: all electron–electron repulsions are either neglected or averaged out through the use of empirical values for the integrals, *i.e.*, empirical parameters. The fight to eliminate this approximation has been a long and tedious one. Probably the first major success was the calculation of benzene by Goeppert-Mayer and Sklar,[2] but their

(1) A. Streitwieser, "Molecular Orbital Theory for Organic Chemists," John Wiley & Sons, Inc., New York, N. Y., 1961.
(2) M. Goeppert-Mayer and A. L. Sklar, *J. Chem. Phys.*, **6**, 645 (1938).

treatment did not become general until the formulation of the SCF method by Roothaan[3] and its application by Pariser and Parr[4] and by Pople.[5] The difficulty here lies in the fact that, for the calculation of electron–electron interactions, one needs to know the orbital of the electron in order to calculate its interaction with another electron. The problem is solved by assuming a set of orbitals, which are used to construct a Hartree–Fock matrix. From this matrix a new set of orbitals and orbital energies are calculated, which in turn are used to construct a new Hartree–Fock matrix. This process is repeated until the new orbital set is identical with the previous one. At this point the results are self-consistent, and hence the name *self-consistent field* (SCF) calculations. The methods of Pariser, Parr, and Pople (PPP) have become the standard methods of π-electron calculation.

Next, chemists were not satisfied with the restriction to planar molecules and to the π-electron systems. Many attempts to calculate nonplanar molecules with the HMO and PPP methods were apparently successful, but it was soon recognized that such success was restricted to very limited series of molecules. Consequently methods were sought to permit inclusion of all valence electrons in the calculations.

One of the earliest approaches to a treatment of all valence electrons is due to Wolfsberg and Helmholz,[6] who calculated the electronic structure and spectra of MnO_4^-, CrO_4^{2-}, and ClO_4^-. Two types of integrals are needed in these calculations; the diagonal matrix elements, H_{ii}, were taken from valence-state ionization potentials, with some arbitrary adjustments, and the off-diagonal H_{ij} were taken from an approximation suggested by Mulliken, $H_{ij} = kS_{ij}(H_{ii} + H_{jj})/2$. This approximation has been used by many authors, and quite a number of calculations have been made by this method and various modifications thereof.[7] However, the most widely accepted, and apparently most generally useful, is the extended Hückel method (EHMO) of Hoffmann,[8] who has introduced a systematic parametrization into an all-valence-electron Hückel-type approximation.

Further progress in all-valence-electron calculations required the development of methodology for the introduction of, at least, the most important electron repulsion integrals, and consequently SCF methods. Early attempts of making such calculations were made by Pohl and coworkers[9] and by Klopman,[10] but no general methodology was developed until about 1965.

The present paper will attempt to summarize the state of these attempts, and, it is hoped, will convey to the experimental chemist a feeling and knowledge for what methods are available, and what type of results may be expected from them.

Methods

SCF Theory. In order to discuss intelligently the various methods proposed for all-valence-electron semiempirical SCF calculations in the past few years, it appears most profitable to first outline a general SCF MO method and focus on the various approximations involved and on all the integrals required.[11] It is particularly in the area of the semiempirical choices for integrals that the various presently available methods differ.

All quantum mechanical methods concerned with stationary—*i.e.*, essentially time-independent properties—start with the time-independent Schrödinger equation

$$H\Psi_i = E_i\Psi_i \qquad (1)$$

Here H is the total Hamiltonian operator of the system of particles; Ψ_i, which is a function of all space and spin coordinates of all particles in the system, is the wave function describing the i stationary state; and E_i is the corresponding eigenvalue, *i.e.*, the energy associated with that state.

After application of the Born–Oppenheimer approximation, which factors Ψ into a nuclear and an electronic part, eq 1 can be interpreted as involving only the electronic coordinates, and involving the internuclear distances only as constants (parameters). For most purposes, the Hamiltonian operator can be considered as consisting of a kinetic energy term for each electron, of the form $-1/2\nabla_i^2$, and the various nuclear–nuclear, nuclear–electron, and electron–electron potential energy terms. These are, respectively[12]

$$H = \sum_{\mu} (-1/2\nabla_\mu^2) + \sum_{M<N} \frac{Z_M Z_N}{r_{MN}} - \sum_{M\nu} \frac{Z_M}{r_{M\nu}} + \sum_{\mu<\nu} \frac{1}{r_{\mu\nu}}$$

$$(2)$$

where Z_M, Z_N are the charges on the Mth and Nth nuclei, and ν and μ are the indices for the electrons. The summations over all μ (and ν) extend over all electrons, those over M (and N) over all nuclei, and $\sum_{\nu<\mu}$ signifies a double summation over all μ, and over all $\nu < \mu$. The function Ψ of eq 1 now is a function of all the electronic space and spin coordinates only. Only if one is interested in magnetic properties need one consider a series of additional magnetic terms.

In principle, eq 1 represents a complete formulation of the problem. Unfortunately, however, it is practically impossible to find a solution for Ψ satisfying this

(3) C. C. J. Roothaan, *Rev. Mod. Phys.*, **23**, 69 (1957).

(4) R. Pariser and R. G. Parr, *J. Chem. Phys.*, **21**, 466 (1953).

(5) J. A. Pople, *Trans. Faraday Soc.*, **49**, 1375 (1953); *J. Phys. Chem.*, **61**, 6 (1957).

(6) M. Wolfsberg and L. Helmholz, *J. Chem. Phys.*, **20**, 837 (1952).

(7) See, *e.g.*, G. J. Ballhausen and H. B. Gray, *Inorg. Chem.*, **1**, 111 (1962); L. C. Cusachs, *et al.*, *J. Chem. Phys.*, **43**, S157, S160 (1965); *J. Phys. Chem.*, **71**, 1060 (1967); S. P. McGlynn, *et al.*, *J. Chem. Phys.*, **44**, 1865 (1966), and following papers.

(8) R. Hoffmann, *ibid.*, **39**, 1397 (1964); **40**, 3247, 2474, 2480 (1964).

(9) H. A. Pohl, R. Rein, and K. Appel, *ibid.*, **41**, 3385 (1964); F. E. Harris and H. A. Pohl, *ibid.*, **42**, 3648 (1965).

(10) G. Klopman, *J. Am. Chem. Soc.*, **86**, 1463, 4550 (1964); **87**, 3300 (1965).

(11) R. G. Parr, "Quantum Theory of Molecular Electronic Structure," W. A. Benjamin, Inc., New York, N. Y., 1964.

(12) For the sake of convenience, all quantities are expressed here in atomic units.

equation. Consequently we are reduced to making a long series of approximations.

The first approximation stems from the realization that the solution of eq 1 would be greatly simplified if we could factor the many electron functions $\Psi(1, 2 \ldots)$ into a series of one-electron functions, $\psi(1), \psi(2) \ldots$, where each $\psi(j)$ is a one-electron function dependent only on the space and spin coordinates of electron j, a spin orbital. Unfortunately this product implies that we can tell which electron occupies which spin orbital, an obvious impossibility. Application of the Pauli principle in its most general form, that any wave function must be antisymmetric with respect to exchange of any two electrons, $i.e.$, that any wave function must change sign when the coordinates of any two electrons are exchanged, leads to the formulation of Ψ as a Slater determinant which is the determinant of all the possible permutations of the electronic coordinates among all the spin orbitals, each permutation with its proper sign.

$$\Psi = |\psi_1(1)\psi_2(2)\cdots\psi_n(n)| \tag{3}$$

The next approximation usually made in molecular quantum mechanics is to express each of the spin orbitals of eq 3 as a linear combination of atomic orbitals, the basis set, multiplied by a spin function, σ.

$$\psi_i = \left(\sum_\nu c_{i\nu}\phi_\nu\right)\sigma_i \tag{4}$$

σ may be either of the two spin functions α or β. The ϕ_ν are the various atomic orbitals of the various atoms of the molecule. The HMO and PPP methods are characterized by taking into account only a single atomic orbital, the p_π orbital of each atom, but in the methods under consideration here the basis set extends over a larger set of orbitals, usually one for H, four for first-row atoms, and nine for transition metals.

With these approximations introduced, finding the proper Ψ for eq 1 becomes equivalent to finding the proper set of $c_{i\nu}$ in eq 4. At this point, we restrict our development to closed shells, $i.e.$, to wave functions in which each MO is occupied by two electrons. According to the variational principle, this set of $c_{i\nu}$ is found by minimizing E_i of eq 1 with respect to all the $c_{i\nu}$. The orbital minimization generates a set of linear homogeneous equations, the solution of which requires that the associated secular determinant vanishes.

$$|F_{\mu\nu} - S_{\mu\nu}\epsilon_i| = 0 \tag{5}$$

Further problems arise from the need to evaluate the elements of the matrix corresponding to this determinant, in particular the $F_{\mu\nu}$, the elements of the Fock matrix in the language of Roothaan.[3]

The $F_{\mu\nu}$ are frequently split into two parts (eq 6),

$$F_{\mu\nu} = H_{\mu\nu} + G_{\mu\nu} \tag{6}$$

the second of which, $G_{\mu\nu}$, is dependent on the occupation of all orbitals, whereas the first, $H_{\mu\nu}$, is independent of occupation number. These integrals can be expressed as

$$H_{\mu\nu} = \int \phi_\mu[-\tfrac{1}{2}\nabla^2 - \sum_M V_M]\phi_\nu \, d\tau \tag{7}$$

$$G_{\mu\nu} = \sum_{\lambda\sigma} p_{\lambda\sigma}[\langle\mu\nu|\lambda\sigma\rangle - \tfrac{1}{2}\langle\mu\sigma|\nu\lambda\rangle] \tag{8}$$

where

$$\langle\mu\nu|\lambda\sigma\rangle = \iint \phi_\mu(1)\phi_\lambda(2) \frac{1}{r_{12}} \phi_\nu(1)\phi_\sigma(2) \, d\tau_1 d\tau_2 \tag{9}$$

and the $p_{\lambda\sigma}$ are the elements of the bond order matrix

$$p_{\lambda\sigma} = 2 \sum_i^{occ} c_{i\lambda}c_{i\sigma} \tag{10}$$

with the summation extending over all occupied orbitals.

Semiempirical Integral Evaluation. In $a\ priori$ methods, all of the integrals of eq 7–9 are evaluated rigorously by introducing some particular functional form for the basis functions ϕ occurring in eq 4. Also V_M is replaced by its correct value, Z_M/r, where Z_M is the charge of nucleus M and r is the distance from this nucleus to the electron. In the semiempirical methods used here this substitution is not possible since V_M refers to the interaction of the electron with the atomic core, $i.e.$, the nucleus and inner electrons. Furthermore, the semiempirical methods estimate at least some of the integrals from empirical data; at the same time, in order to hold the number of integrals to be estimated within reason, and to reduce computational labor, some integrals are neglected, and it is hoped that the empirical choices for others are able to compensate for this neglect. The various methods which have recently been proposed differ basically only in these integral choices and approximations. Pople and coworkers have proposed three different levels of approximation, and we shall outline these in later sections and then compare the other methods with those of Pople.

At this point it may be worthwhile to make a distinction between, at least, two basically different types of semiempirical parameters. In the Hückel method, all basic integrals are treated as disposable parameters, for which values are chosen in order to make the results of calculations reproduce the observed quantities of interest in one or a few reference compounds. The adequacy of these choices is then tested by making calculations for further molecules and comparing calculated and experimental results. Parameters of this sort may be called adjustable or disposable empirical parameters, and it requires quite a number of successful calculations to establish their validity.

In going to more elaborate methods, one might have been tempted to continue choosing all integrals in the same fashion. However, the procedure is cumbersome, and not very satisfying, and requires an ever increasing body of experimental data for comparison to establish validity. Consequently workers in semiempirical methods have been increasingly using either completely theoretically calculated integrals (as, $e.g.$, the Γ_{MN} in the CNDO method) on integrals obtained from a the-

oretical treatment of experimental data on simple systems, primarily atomic spectroscopic data, as, e.g., the $\Gamma_{\mu\mu}$ of the Pariser–Parr–Pople method. Although such parameters still must be classed as semiempirical, they are no longer adjustable, and their choice is made essentially before calculations are made. The methods to be discussed here will make use of both types of parameters.

The CNDO Method. The CNDO (complete neglect of differential overlap) method[13] makes the approximation that the product $\phi_\mu\phi_\nu$ of any pair of orbitals ϕ_μ and ϕ_ν is always zero everywhere, as long as ϕ_μ and ϕ_ν are not different centers. This approximation is quite analogous to one made in the PPP method, but each atom or center (M, N, etc.) contributes several orbitals (denoted by μ, ν, etc.) to the system. This approximation leads to two results: (1) the overlap integral, $S_{\mu\nu}$, between any pair of orbitals is zero; (2) all integrals $\langle\mu\nu|\lambda\sigma\rangle$ are zero unless $\mu = \nu$ and $\lambda = \sigma$. The nonzero integrals $\langle\mu\mu|\lambda\lambda\rangle$ are called $\Gamma_{\mu\lambda}$. If the results of such a calculation are to be independent of the arbitrary assignment of coordinate axes to the molecule, this approximation requires an additional approximation, that $\Gamma_{\mu\lambda}$ be independent of the particular nature of the orbitals ϕ_μ and ϕ_λ and depend only on the centers M and N on which ϕ_μ and ϕ_λ are located; thus all the various $\Gamma_{\mu\lambda}$ between orbitals of M and N are approximated by a single Γ_{MN}.

In the original CNDO calculations, the Γ_{MN} are approximated by the purely theoretical values of the integral $\langle s_M s_M | s_N s_N \rangle$ of the valence shell s orbitals.

The diagonal element $H_{\mu\mu}$ may be split into two parts, one essentially an atomic term of the atom M on which ϕ_μ is centered, and the other the interaction of the electron in ϕ_μ with the cores of all other atoms.

$$H_{\mu\mu} = \langle\phi_\mu|-^1/_2\nabla^2 - V_M|\phi_\mu\rangle - \sum_{M\neq N}\langle\phi_\mu|V_N|\phi_\mu\rangle$$

$$= U_\mu - V_{MN}$$

The first term is approximated from atomic spectroscopy. With the differential overlap approximation made above, all electronic states of a given configuration have the same energy

$$E(X,2s^m2p^n) = mU_{2s} + nU_{2p} +$$

$$^1/_2(m + n)(m + n - 1)\Gamma_{xx} + C$$

where C represents terms coming from the inner electrons. This energy corresponds to a properly weighted mean of all electronic states of this configuration. To evaluate a single U_μ one takes the proper ionization process, as, e.g.

$$I_s(X,2s^m2p^n) = E(X^+,2s^{m-1}2p^n) - E(X,2s^m2p^n)$$

$$= -U_{2s} - (m + n - 1)\Gamma_{xx}$$

$$U_{2s} = -I_s(X,2s^m2p^n) - (m + n - 1)\Gamma_{xx}$$

In later papers,[14] the integral U_μ was modified by being taken as the average of the ionization potential (I) and the electron affinity (A).

$$-A_\mu = U_\mu + Z_M\Gamma_{MM}$$

$$U_\mu = ^1/_2(I_\mu + A_\mu) - (Z_M - ^1/_2)\Gamma_{MM}$$

The second term in $H_{\mu\mu}$, V_{MN}, was approximated in the first CNDO work by the integral

$$\int s_M^2 \frac{Z_N}{r_N}\,d\tau$$

but experience showed that it was better to approximate V_{MN} by $Z_N\Gamma_{MN}$, where Z_N is the charge of the core of N. This difference in approximation is the second difference between CNDO/1 and CNDO/2, as used by Pople and coworkers.[13,14]

Finally, the off-diagonal elements $H_{\mu\nu}$ are taken as zero if ϕ_μ and ϕ_ν are centered on the same atom. If ϕ_μ and ϕ_ν are centered on different atoms M and N

$$H_{\mu\nu} = ^1/_2(\beta_M^0 + \beta_N^0)S_{\mu\nu}$$

The β_M^0 are treated as adjustable empirical parameters and chosen to give the best agreement between CNDO and a priori LCAO SCF calculations on diatomic molecules.

Thus, the matrix elements of the Fock matrix in the CNDO/2 formulation are

$$F_{\mu\mu} = -^1/_2(I_\mu + A_\mu) + [(p_{MM} - Z_M) - ^1/_2(p_{\mu\mu} - 1)] \times$$

$$\Gamma_{MM} + \sum_{N\neq M}(p_{NN} - Z_N)\Gamma_{MN}$$

$$F_{\mu\nu} = ^1/_2 S_{\mu\nu}(\beta_M^0 + \beta_N^0) - ^1/_2 p_{\mu\nu}\Gamma_{MN}$$

$$\Gamma_{MN} = F^0 = \iint s_M(1)s_N(2)\frac{1}{r_{12}}s_M(1)s_N(2)\,d\tau_1 d\tau_2$$

In these expressions $p_{\mu\mu}$ refers to the diagonal elements of the bond order matrix, i.e., the electron density in a given orbital, while p_{MM} (or p_{NN}) is the sum over the diagonal elements of the bond order matrix for all orbitals belonging to atom M (or N).

Other Methods. In the INDO (intermediate neglect of differential overlap) method,[15] the only major change in approximations is that the differential overlap is not neglected in integrals $\langle\mu\nu|\lambda\sigma\rangle$ if all orbitals ϕ_μ, ϕ_ν, ϕ_λ, and ϕ_σ are centered on the same atom, i.e., in one-center integrals. For a basis set of pure s, p, and d orbitals, the only integrals arising are of the form $\langle\mu\mu|\nu\nu\rangle$ and $\langle\mu\nu|\mu\nu\rangle$, and these were expanded in terms of the Slater–Condon parameters F^0, G^1, and F^2. F^0 is identical with Γ_{MM}, and G^1 and F^2 are taken from Slater's compilation of values giving best fits with atomic spectroscopic values.

The I_μ terms are also modified by introduction of the above integrals; the forms vary for the different atoms, and the results can be taken from the original paper.

(13) (a) J. A. Pople, D. P. Satry, and G. A. Segal, J. Chem. Phys., 43, S129 (1965); (b) J. A. Pople and G. A. Segal, ibid., 43, S136 1965).

(14) J. A. Pople and G. A. Segal, ibid., 44, 3289 (1966).
(15) J. A. Pople, D. L. Beveridge, and P. A. Dobosh, ibid., 47, 2026 (1967).

Dixon[16] has independently proposed a modification, EMZDO (exchange-modified zero differential overlap), of the CNDO/1 method which achieves the same purposes and is almost identical with the INDO treatment. The term $V_{M\nu}$ is set equal to Γ_{MN}. One-center exchange integrals $\langle \mu\nu|\mu\nu \rangle$ are not neglected, and it must be assumed that theoretical values are used. One-center Coulomb integrals are not, as in the CNDO methods, made independent of orbital type, and again it must be assumed that theoretical values are used.

The final approximation introduced by Pople and coworkers is the NDDO (neglect of diatomic differential overlap),[13] in which all integrals $\langle \mu\nu|\lambda\sigma \rangle$ in which ϕ_μ and ϕ_ν are centered on one atom, M, ϕ_λ and ϕ_σ on one atom which may or may not be M, are retained in the calculation. This method requires a very large number of such integrals, and no approximations have been proposed by Pople and coworkers, nor have any calculations by this method been reported.

The method of Dewar and Klopman,[17] based on earlier attempts by Dewar and coworkers[18] and by Klopman,[7b] turns out to be very similar to that of Pople and coworkers, and is described by the authors as PNDDO (partial neglect of diatomic differential overlap). In this approach, the basis set on each pair of atoms is transformed to a common coordinate system for the evaluation of the two-center integrals, then all three and four atom integrals are neglected. However, for each pair of unequal atoms, ten different integrals are required, of the forms

$$\langle s_M s_M | s_N s_N \rangle, \quad \langle s_M s_M | \sigma_N \sigma_N \rangle, \quad \langle s_M M | \pi_N \pi_N \rangle,$$

$$\langle \sigma_M \sigma_N | \sigma_N \sigma_N \rangle, \quad \langle \sigma_M \sigma_M | \pi_N \pi_N \rangle, \quad \langle \pi_M \pi_M | \pi_N \pi_N \rangle,$$

$$\langle \pi_M \pi_M | \pi'_N \pi'_N \rangle$$

Thus the problem of choosing semiempirical expressions or values for integrals is multiplied about tenfold. The various integrals are grouped in three groups, (a) those involving p_σ orbitals for both electrons, (b) those involving them for only one electron, and (c) integrals not involving p_σ orbitals. Three separate formulas are used, respectively, for the three groups

$$\langle \mu\mu|\nu\nu \rangle = [r_{\mu\nu}^2 + (\rho_\mu + \rho_\nu)^2 T_{\mu\nu}^2]^{-1/2} \qquad (a)$$

$$\langle \mu\mu|\nu\nu \rangle = [r_{\mu\nu}^2 + (\rho_\mu + \rho_\nu T_{\mu\nu})^2]^{-1/2} \qquad (b)$$

where ν is the ρ_σ orbital, and

$$\langle \mu\mu|\nu\nu \rangle = [r_{\mu\nu}^2 + (\rho_\mu + \rho_\nu)^2]^{-1/2} \qquad (c)$$

with $T_{\mu\nu} = \exp[-r_{\mu\nu}/2(\rho_\mu + \rho_\nu)]$. ρ_μ and ρ_ν are derived from spectroscopic data, and used to ensure that the integrals converge on the proper one-center integrals at $r_{\mu\nu} = 0$. The ρ are defined by

$$\rho_\nu = e^2/2A_M^-$$

where A_M^- is the average repulsion of any two valence electrons of opposite spin in any two atomic orbitals of atom N on which ϕ_ν is centered.[7b] Thus, the method has reduced the integrals to three types, based on the same spectroscopic data and calculated by slightly modified formulas, presumably giving expression to distinctive behavior of p_σ orbitals.

The one-center integrals $\langle \mu\mu|\nu\nu \rangle$ are approximated as A_M^- (ϕ_μ and ϕ_ν both are atomic orbitals of M). The one-center exchange integral $\langle \mu\nu|\mu\nu \rangle$ is approximated as $A_M^- - A_M^+$, where A_M^+ is the average repulsion of any two valence electrons of like spin on atom M, independent of the orbitals they occupy. The ionization potentials needed for the diagonal elements of the Fock matrix are evaluated from atomic spectroscopic data, in a manner not unlike that of Pople and coworkers.

Finally, the electron repulsion independent term $H_{\mu\nu}$ of the Fock matrix is evaluated by

$$\beta_{\mu\nu} = \beta_M^0 \beta_N^0 S_{\mu\nu} (I_\mu + I_\nu) [r_{\mu\nu}^2 + (\rho_\mu + \rho_\nu)]^{-1/2}$$

where ϕ_μ is centered on M, ϕ_ν on N.

I_μ and I_ν are the proper valence-state ionization potentials; the β^0 are atomic, not orbital parameters. The β_M^0 were fitted empirically. Since these calculations were designed to produce heats of formation, ΔH_f, it was essential to include good approximations for the core repulsions; these were evaluated as

$$C_{MN} = E_{MN} + [Z_M Z_N(1/r_{MN} - E_{MN})] \times$$

$$\exp(-\alpha_{MN} r_{MN})$$

where the A_M are the core charges, r_{MN} is the distance between atoms M and N, E_{MN} is the electronic repulsion between the neutral atoms

$$E_{MN} = \sum_\mu^M \sum_\nu^N \Gamma_{\mu\nu}$$

α_{MN} is given by

$$\alpha_{MN} = \sqrt{\alpha_{MM} \alpha_{NN}}$$

and α_{MM} is treated as an adjustable parameter.

A third distinct method was introduced by Del Bene and Jaffé[19] with the intent of calculation of absorption spectra. The method used was formally exactly analogous to the CNDO/2 formulation, but with the following modification in the parametrization. Γ_{MM} was approximated as in the Pariser–Parr–Pople method by $I_M - A_M$.

The $H_{\mu\nu}$ terms were evaluated as

$$H_{\mu\nu} = \frac{1}{2}\kappa(\beta_M^0 + \beta_N^0)S_{\mu\nu}$$

β_M^0 are purely empirical atomic parameters newly chosen and apparently not related to Pople's values.[13,14] $\kappa = 1$ when $S_{\mu\nu}$ measures σ-type overlap, $\kappa = 0.585$, a purely empirical value, when $S_{\mu\nu}$ measures π-type overlap. This distinction requires the same type of transformation made by Dewar and Klopman,[17] but the

(16) R. N. Dixon, *Mol. Phys.*, **12**, 83 (1967).
(17) M. J. S. Dewar and G. Klopman, *J. Am. Chem. Soc.*, **89**, 3089 (1967).
(18) M. J. S. Dewar, G. J. Gleicher, and C. C. Thompson Jr., *ibid.*, **88**, 1349 (1966), and previous papers.

(19) J. Del Bene and H. H. Jaffé, *J. Chem. Phys.*, **48**, 1807 (1968).

process is automatically carried out in the electronic computation.

Invariance Problems

Pople and coworkers,[13] in the original exposition of the differential overlap methods, required their methods to be invariant to two types of transformation: (1) a unitary transformation between orbitals of the various individual atoms; this corresponds to rotation of local coordinate systems of atoms or the total coordinate system for the molecule; let us call the invariance to this process *rotational invariance;* (2) a unitary transformation between s and p orbitals; this process corresponds to transition between a basis set made up of pure atomic orbitals and one composed of hybrid orbitals; let us call invariance to this process *hybridization invariance.*

Pople's requirement that results of CNDO calculations be both rotationally and hybridizationally invariant led to a number of approximations, in particular the use of Γ_{MM} as an atomic property, independent of the properties of the orbitals involved.

Dewar and Klopman's[17] neglect of three- and four-orbital two-center integrals in the PNDDO method violates the rotational invariance criterion in principle, but appears to be practically without importance.

Hybridization invariance appears of importance only if we want to treat a portion of a molecule, making the assumption that certain bonds or groups are well enough known, and need not further be included. Thus, computational labor might be greatly reduced by treating methyl or *t*-butyl groups as a pseudo atom with a single hybrid atomic orbital.

Whether rotational invariance is worth the price in approximations is one of the major arguments of the day. Some authors argue that it must be maintained at all costs. Others counter with the plausible argument that results differ only by terms which are being neglected. This argument, however, is fallacious: terms are neglected not because they are necessarily small; on the contrary, the semiempirical parametrization is designed to correct for an average effect of the omitted terms, and for just these reasons truly theoretical integrals often fail to yield good results. Thus it must be concluded that rotational invariance *is* essential, while hybridization invariance, though desirable for certain types of extension, is not critical.

Results

CNDO. In the relatively short time since these methods have been postulated an amazing amount of work has been done with their use.

Pople and Segal[13b] have tested the CNDO method in the CNDO/1 approximation by calculating dipole moments for a number of diatomic molecules, with moderate success. However, the main purpose of these calculations was to calibrate the $\beta_M{}^0$ so that the CNDO results would give good agreement with complete SCF calculations.

Further calculations on polyatomic molecules gave

again reasonable dipole moments and bending force constants.[13b] For many of the same molecules, bond angles were calculated, and found to agree with experimentally determined values to an amazing degree. All geometric parameters, except those varied, were taken from experimental determinations and held constant. Attempts to calculate bond lengths met with failure.

Upon introduction of the CNDO/2 method, Pople and Segal[14] calculated bond angles, dipole moments, and bending force constants of a large number of AB_2 and AB_3 molecules, largely with very good success. Pople and Gordon[20] have calculated dipole moments of a large number of simple organic compounds containing N, O, and F atoms, in reasonable agreement with experimental values. The charge distributions calculated show widespread alteration of polarity in both saturated and unsaturated systems. Segal and Klein[21] have calculated dipole moment derivatives for some small molecules and have found that the magnitudes, and particularly the signs of the derivatives, are given reasonably well. The main problems arise where derivatives are small because of cancellation of large terms in the calculation, and then even the sign may be in doubt.

After the failure to obtain good bond lengths by the CNDO/1 method, Segal[22] repeated such calculations by the CNDO/2 variation. The results for the bond lengths of some diatomic molecules are amazingly good, but the stretching force constants are consistently too large.

Bloor and Breen[23] have applied the CNDO/2 method to ground-state properties of some larger nitrogen and oxygen heterocycles. Their calculations of dipole moments agree very well with observed results. These authors also find that total charge densities are well related to ^{13}C chemical shifts. However, the ionization potentials of a series of heterocycles are too high by 3.5–5 eV, and, in the azines, correspond to ionization of σ electrons, contrary to experiment.[24a]

Wiberg[24b] has applied the CNDO method (presumably CNDO/2) to a series of hydrocarbons and their cation radicals and to some free radicals and the corresponding carbonium ions. By applying some modifications to the parameters used, he was able to calculate the geometry in good agreement with experiment. Using a semiempirical scaling factor on the energies, he was able to reproduce the heats of atomization and of formation satisfactorily.

Davies[25] has made some calculations on some fluorinated benzenes and nitrobenzenes and on the radical anions of the latter. Dipole moments are reproduced in good approximation, and hyperfine coupling con-

(20) J. A. Pople and M. Gordon, *J. Am. Chem. Soc.*, **89**, 4253 (1967).
(21) G. A. Segal and M. L. Klein, *J. Chem. Phys.*, **47**, 4236 (1967).
(22) G. A. Segal, *ibid.*, **47**, 1876 (1967).
(23) J. E. Bloor and D. L. Breen, *J. Am. Chem. Soc.*, **89**, 6835 (1967).
(24) (a) J. Del Bene and H. H. Jaffé, *J. Chem. Phys.*, **49**, 1221 (1968); (b) K. B. Wiberg, *J. Am. Chem. Soc.*, **90**, 59 (1968).
(25) D. W. Davies, *Mol. Phys.*, **13**, 465 (1967).

stants of the radical anions follow the same general pattern as calculated spin densities.[26] Nmr chemical shifts are calculated in good approximation except for one compound, p-fluoronitrobenzene.

An extension of the CNDO method to second-row elements (Na to Cl) proved less straightforward since 3d orbitals were found to be essential in these calculations.[26] Although bond angles were substantially independent of the 3d orbitals, dipole moments could not be approximated without these orbitals. Moreover, parametrization for the 3d orbitals was difficult; their inclusion in the orbital-independent atomic parametrization was not satisfactory. Hybridization invariance is lost, but results are significantly improved if the d orbitals are treated different from s and p orbitals. The following are proposed.

$$Z(3d) = 0.75 Z(3p)$$

$$\beta^0(3d) = [2U_{3d}/(U_{3s} + U_{3p})]\beta^0(3s,3p)$$

The original CNDO/2 parametrization has been used by Kroto and Santry[27] to calculate the geometry of excited states of a few selected molecules. The calculations are made using a virtual orbital approximation, but the energy is newly minimized, using excited-state bond lengths and varying the angles of interest. Although the excitation energies are consistently too high by 0.5–4 eV, the calculated bond angles in the excited states are again amazingly good. The same authors[28] have introduced an open-shell calculation in an attempt to further improve excited-state calculations. Acetylene, formaldehyde, and HCF are treated in both papers; both bond angles and excitation energies are slightly improved. However, the improvement does not seem to justify the greatly expanded labor.

An attempt to calculate spectra and ionization potentials for methyl-substituted borazines and benzenes may be called less than successful.[29] Another attempt to apply the original CNDO/2 parametrization to the calculation of spectra of large molecules was undertaken by Clark and Ragle.[30] These authors found states arising from transitions involving σ orbitals intermingled with the desired $\pi-\pi^*$ states in ethylene and benzene, and their calculations may be called unsuccessful.[30]

The same observation of intermingling of σ and π states has led Del Bene and Jaffé[19,31] to a complete reparametrization of the CNDO/2 formalism and the addition of a limited configuration interaction calculation. The method, which was outlined in the preceding section and is a virtual orbital approximation, was calibrated to reproduce the spectra of benzene, pyridine, and a few related molecules. To date, five

adjustable empirical parameters have been used, the four $\beta_M{}^0$ for M = H, C, N, and O, and K. These five values have permitted the reproduction of the $\pi-\pi^*$ and n–π^* transitions in a large number of medium size compounds. Throughout, in benzene analogs, the band corresponding to the $^1B_{1u}$ band of benzene, the ^1L_a band (in Platt notation), is calculated at too low an energy. This fault of the calculation is analogous to what has been found in Pariser–Parr–Pople calculations and is undoubtedly due to the use of the Pariser interpolation formula. Replacement of this formula by a Mataga-type formula will probably overcome this difficulty. Relative intensities of absorption bands are well reproduced, and even oscillator strengths are usually reasonably approximated. As far as is known, polarizations are correctly assigned.

In a few small molecules,[32] spectra not involving π-electron systems are well represented. On the other hand, attempts to calculate bond angles were not too successful. In a few heterocyclic molecules, ionization potentials are much better reproduced than by CNDO/2. Nmr correlations are similar to those of Bloor and Breen,[23] but dipole moments are not as well represented as by CNDO/2 parametrization.

The INDO method was introduced by Pople and coworkers[15] largely to be able to deal with spin densities, which are not properly taken care of in the CNDO formalism. The method was first tested by recalculating bond angles for many of the molecules previously treated by CNDO/1 and CNDO/2. Bond angles are little affected, but an improved description of the relative energies of singlet and triplet states is obtained. A very similar approximation,[16] EMDZO (exchange-modified zero differential overlap), leads to similar results. The method leads to some reasonable excitation energies for the lowest excited states of a few small radicals and molecules, CH, CH_2, NH_2, and formaldehyde.

However, the INDO formalism was developed primarily to account for esr hyperfine coupling constants in terms of spin densities. In a massive study, Pople, Beveridge, and Dobosh[33] have calculated spin densities on 57 compounds and correlated the results with isotropic hyperfine coupling constants for protons, ^{13}C, ^{14}N, ^{17}O, and ^{19}F. The slopes of the least-squares lines of the plots of coupling constants vs. spin densities for each type of nucleus are the only empirically adjustable parameters in the entire calculation, and represent the integral

$$\int |\phi_{sM}(r_M)|^2$$

which is the density of the valence s electrons of atom M evaluated at its nucleus and which, at the level of theory used, cannot be evaluated theoretically. The results obtained are excellent.

(26) D. P. Santry and G. A. Segal, J. Chem. Phys., **47**, 158 (1967).

(27) H. W. Kroto and D. P. Santry, ibid., **47**, 792 (1967).

(28) H. W. Kroto and D. P. Santry, ibid., **47**, 2736 (1967).

(29) P. M. Kuznesof and D. F. Shriver, J. Am. Chem. Soc., **90**, 1683 (1968).

(30) P. A. Clark and J. L. Ragle, J. Chem. Phys., **46**, 4235 (1967); P. A. Clark, ibid., **48**, 4795 (1968).

(31) J. Del Bene and H. H. Jaffé, ibid., **48**, 4050 (1968).

(32) J. Del Bene and H. H. Jaffé, ibid., **50**, 563 (1969).

(33) J. A. Pople, D. L. Beveridge, and P. Dobosh, J. Am. Chem. Soc., **90**, 4201 (1968).

In a very elegant study, Beveridge and Miller[34] have analyzed isotope effects on isotropic hyperfine coupling constants by calculation of vibronic effects in CH_3 and CD_3 with INDO wave functions. The trends of the coupling constants are very well reflected by these calculations.

PNDDO. The most elaborate calculation reported within the NDO (neglect of differential overlap) framework is Dewar and Klopman's[17] calculation of absolute heats of formation, ΔH_f, of a long series of hydrocarbons. The calculations contain only four adjustable empirical parameters, the β_M^0 for H and C and the core repulsion parameters α_H and α_C. For all saturated hydrocarbons, and for singly unsaturated and conjugated hydrocarbons, deviations in ΔH_f do not exceed 5 kcal/mol, with an average deviation of 2.3 kcal for 17 compounds. Only in allene and the acetylenes are the deviations larger, of the order of 1 eV. In most cases the difference between conformational isomers is also calculated to within 2–3 kcal/mol, and, with the exception of the normal–branched hydrocarbon pairs, the most stable species has the lowest energy.

Dewar and Klopman[17] also compare their orbital energies with series of ionization potentials obtained from photoionization spectra. The correspondence is quite good, with an average deviation of 0.7 eV. The first ionization potential is too high by about 1 eV, except for ethylene and acetylene; all values for acetylene are low, again showing that the calculations do not well reflect this compound, or triple (and presumably cumulative double) bonds in general. All observed ionization potentials except one (in ethylene) have calculated counterparts, but some calculated ones are not observed. Finally, small dipole moments are calculated for some saturated hydrocarbons, in reasonably good agreement with experiment.

This work was extended by Baird and Dewar[35] to a series of cyclopropanes and cyclopropenes. They were able to calculate strain energies in reasonable approxi-mation compared with values obtained from bond energy additivity relations.

Conclusions

From the results reported here it may be concluded that the various NDO methods discussed are surprisingly successful in allowing reasonable calculations on molecules of varying complexity. However, it appears that the same approximations or parametrizations are not equally successful in the calculation of different properties.

(1) CNDO/2 calculations have been found quite satisfactory in reproducing molecular geometry, and even bond lengths in small molecules. They are less satisfactory in the calculation of dipole moments and are apparently incapable of reproducing heats of formation, ionization potentials, spectra, or spin densities.

(2) CNDO calculations with a special parametrization for spectra are very successful in reproducing electronic spectra, both band positions and intensities. The same calculations give better approximations to ionization potentials than the original parameters.

(3) INDO calculations are extremely successful in the calculation of spin densities, and hence of hyperfine coupling constants. At the same time, they give geometries as good, if not better than, CNDO/2, and the dipole moments are improved.

(4) PNDDO calculations have provided heats of formation for a series of hydrocarbons with good accuracy and have given ionization potentials within ~1 eV. However, since no atoms other than C and H are treated, and since results for acetylenes and cumulenes have not been satisfactory, more work remains to be done in this area.

(34) D. L. Beveridge and K. Miller, *Mol. Phys.*, **14**, 401 (1968).
(35) N. C. Baird and M. J. S. Dewar, *J. Am. Chem. Soc.*, **89**, 3966 (1967).