Molecular orbital methods in organic chemistry

John A. Pople

Acc. Chem. Res., **1970**, 3 (7), 217-223 • DOI: 10.1021/ar50031a002

More About This Article

The permalink http://dx.doi.org/10.1021/ar50031a002 provides access to:

- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

Molecular Orbital Methods in Organic Chemistry

JOHN A. POPLE*

Department of Chemistry, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

Received March 23, 1970

It is now more than 40 years since quantum mechanics first provided a complete underlying theory for chemistry and promised to make it a truly mathematical science. Since that time, fulfillment of this promise has often been delayed by mathematical difficulties, and successful quantitative applications of theoretical methods have usually been limited to diatomics and other small molecules. Nevertheless, there has been progress and, during the last few years, it has become possible to use a number of fairly sophisticated molecular orbital methods to study the electronic structure of moderately large molecules. The field is an active one, with many participating research groups, and it seems likely that quantitative theories of electronic structure will play an increasing role in many branches of chemistry in the near future. The aim of this Account is to give a partial survey of some of the methods that have been developed in our laboratory, together with some illustrations of their application. Most of our work has been concerned with organic molecules, where there are only a small number of electrons per atom, but the principles are, of course, more general.

One of the main objectives of a theoretical chemist should be to provide a critical quantitative background for simple qualitative hypotheses about electronic structure. In order to be useful as an independent study, a quantum mechanical theory must be formulated in an unbiased manner, so that no preconceived ideas derived from conventional qualitative discussions are implicitly built in. Molecular orbital theories satisfy this type of condition insofar as each electron is treated as being free to move in a path covering the entire molecular framework.

A further requirement for a useful theory is that the approximate wave functions used must be amenable to detailed interpretation and comparison among related molecules. For example, it is helpful if the calculated electron charge distribution can be easily and realistically divided into contributions of individual atoms which may then be compared with qualitative discussions.

Another long-range objective is the development of quantitative methods to a point at which they can be useful in predicting the geometry and energies of species which are not amenable to direct experimental study. In approaching this objective, we should recognize that there is little prospect of calculating total energies with high absolute accuracy for large molecules. A more modest, but realistic, approach is to adopt several clearly defined levels of approximation which do not represent the ultimate possible for small molecules but

which are simple enough to be widely applied. Each such method is then really a model system implying an entire chemistry of its own. Each can then be tested against real chemistry where experimental data are available and, if there is consistent success, some confidence is acquired in its predictive power. Much recent research has proceeded in this way, both with semiempirical theories involving parametrization with some experimental data and with *ab initio* calculations where one particular method of calculation has been used for a range of different molecules and configurations. We have been concerned with the development of both types of theories at a level simple enough for widespread application to organic molecules.

General Molecular Orbital Theory

Most molecular orbital theories assign electrons in pairs to orbitals ψ_i which are written as linear combinations of a set of basis functions ϕ_μ (eq 1). The ϕ_μ are

$$\psi_i = \sum_\mu c_{\mu i}\phi_\mu \qquad (1)$$

normally centered on the atoms, so expansion 1 is often described as the linear combination of atomic orbital (LCAO) approximation.

It is clear that, as a general rule, more accurate molecular orbitals can be obtained from large basis sets ϕ_μ which permit increased flexibility in the representation. However, this increases the complexity of the calculations and frequently limits applications to small molecules. As a general rule, molecular orbitals are simplest to apply and interpret if the basis set is *minimal*, that is, consists only of the least number of atomic orbitals (of appropriate symmetry) for the atomic ground state. Minimal basis functions for atoms to neon are listed in Table I. If a larger number of ϕ functions are used, the basis set is usually described as *extended*.

Once the molecular orbitals are determined, the charge density can be analyzed in terms of the basis functions ϕ_μ. One major aim of such an analysis is to divide the total charge into parts associated with each such function. If there are two electrons per molecular orbital, the total charge density is

$$\rho = 2\sum_i^{occ}\psi_i^2 = \sum_\mu\sum_\nu P_{\mu\nu}\phi_\mu\phi_\nu \qquad (2)$$

Table I
Minimal Basis Functions

Atoms	Functions
H, He	1s
Li, Be	1s, 2s
B, C, N, O, F, Ne	1s, 2s, $2p_x$, $2p_y$, $2p_z$

where

$$P_{\mu\nu} = 2\sum_i^{occ} c_{\mu i} c_\nu \qquad (3)$$

$P_{\mu\nu}$ is the *density matrix* and contains the detailed information about charge distribution that is implicit in the molecular orbital wave function. The diagonal element $P_{\mu\mu}$ is the coefficient of the distribution $\phi_\mu{}^2$ and is a measure of an electron population for this orbital. The off-diagonal elements $P_{\mu\nu}$ are overlap populations, related to the charge density associated with the overlap functions $\phi_\mu\phi_\nu$. Since this distribution is in the bonding region between ϕ_μ and ϕ_ν, the $P_{\mu\nu}$ elements are sometimes described as bond orders.

A full or *gross population* associated with an atomic orbital was introduced by Mulliken[1] as eq 4, where $S_{\mu\nu}$

$$q_\mu = P_{\mu\mu} + \sum_{\nu(\neq\mu)} P_{\mu\nu} S_{\mu\nu} \qquad (4)$$

is the overlap intergral (eq 5). The gross populations

$$S_{\mu\nu} = \int \phi_\mu \phi_\nu \, d\tau \qquad (5)$$

q_μ add up to the total number of electrons in the molecule and are convenient measures of detailed charge distribution.

The general problem of deriving the LCAO coefficients $c_{\mu i}$ by the variational method was solved by Roothaan[2] who derived eq 6, where the ϵ_i are one-electron energies

$$\sum_\nu (F_{\mu\nu} - \epsilon_i S_{\mu\nu}) c_{\nu i} = 0 \qquad (6)$$

and $F_{\mu\nu}$ is the Fock matrix (eq 7). Here $H_{\mu\nu}$ is the

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma}[(\mu\nu|\lambda\sigma) - {}^1/_2(\mu\lambda|\nu\sigma)] \qquad (7)$$

matrix of the one-electron Hamiltonian for motion in the field of bare nuclei and $(\mu\nu|\lambda\sigma)$ is the two-electron integral 8. Since the density matrix $P_{\mu\nu}$ depends on

$$(\mu\nu|\lambda\sigma) = \iint \phi_\mu(1)\phi_\nu(1)(1/r_{12})\phi_\lambda(2)\phi_\sigma(2) \, dt_1 dt_2 \qquad (8)$$

the molecular orbital coefficients $c_{\mu i}$, eq 4 is not linear and has to be solved by an iterative procedure. It is frequently described as a self-consistent (LCAO–SCF) equation.

The most difficult part of LCAO–SCF theory is usually the evaluation of the large number of two-electron integrals (eq 9). Semiempirical methods usually treat these in a simplified manner and avoid the difficulties. *Ab initio* methods evaluate the integrals but have to use basis functions ϕ_μ for which such integration is possible.

The LCAO–SCF molecular orbital theory described here does, of course, use only a single determinant wave function and takes no account of the correlation between the motions of electrons with antiparallel spins. To proceed beyond this step, it is necessary to use many determinants (configuration interaction). It is well known that such refinement is necessary to give an

adequate description of total binding energies. However, it is much more difficult to apply configuration interaction methods systematically to large molecules, so most research in the theory of the structure of organic molecules, including everything in this paper, is limited to the single-determinant molecular orbital theory. It seems wise to study this level of theory thoroughly as a first stage, finding how it will perform, before attempting a fuller treatment.

Semiempirical Methods

The various semiempirical methods which have been applied to organic molecules start with the LCAO–SCF equations (6) and simplify them in various ways, introducing semiempirical parameters in the process.

The simplest general semiempirical methods are the extended Hückel theories which are developments of the original π-electron MO theory due to Hückel.[3] These theories suppose that the Fock matrix elements for valence orbitals can be approximately treated as independent of the electron distribution. Equation 6 then becomes linear and is easily solved. A common form of the theory, due originally to Mulliken[4] and extensively developed by Hoffmann,[5] uses eq 9, where

$$F_{\mu\mu} = -I_\mu$$
$$F_{\mu\nu} = -{}^1/_2(I_\mu + I_\nu)KS_{\mu\nu} \qquad (9)$$

I_μ is an orbital ionization potential (measuring relative electron-attracting powers of orbitals) and K is a constant.

The principal limitation of extended Hückel theory is its failure to allow for the dependence of the electron-attracting power of an atomic orbital on the number of electrons already on the same and surrounding atoms. It is likely to be most successful in obtaining molecular orbitals for nonpolar molecules where the local net charge on an atom is close to zero. The same approximations also lead to the omission of interactions between polar and ionic groups unless the theory is appropriately modified.

The simplest semiempirical treatment which attempts to take broad account of the main features of electron–electron interaction is the complete neglect of differential overlap (CNDO) method.[6] This approximate method (which treats only valence electrons explicitly) neglects the smaller two-electron integrals (eq 8) which involve the overlap between different ϕ functions and treats the other integrals in a simplified manner. With some further approximations, the matrix elements for the Fock Hamiltonian become eq 10. In these equations, $-{}^1/_2(I_\mu + A_\mu)$ is the Mulliken electronegativity[7] of the orbital ϕ_μ, obtained as the average of the ionization potential and electron affinity. Z_A is the nuclear charge on atom A, so that $q_A - Z_A$ is the net (negative)

(1) R. S. Mulliken, *J. Chem. Phys.*, **23**, 1833, 1841, 2338, 2343 (1955).
(2) C. C. J. Roothaan, *Rev. Mod. Phys.*, **23**, 69 (1951).
(3) E. Hückel, *Z. Physik*, **70**, 204 (1931).
(4) R. S. Mulliken, *J. Chem. Phys.*, **46**, 497, 675 (1949).
(5) R. Hoffmann, *ibid.*, **39**, 1397 (1963).
(6) J. A. Pople, D. P. Santry, and G. A. Segal, *ibid.*, **43**, S129 (1965).
(7) R. S. Mulliken, *ibid.*, **2**, 782 (1934).

$$F_{\mu\mu} = -\tfrac{1}{2}(I_\mu + A_\mu) + [(q_A - Z_A) -$$
$$\tfrac{1}{2}(q_\mu - 1)]\gamma_{AA} + \sum_{B(\neq A)} (q_B - Z_B)\gamma_{AB}$$

$$(\phi_\mu \text{ on atom A}) \quad (10a)$$

$$F_{\mu\nu} = \beta^\circ_{AB} S_{\mu\nu} - \tfrac{1}{2}P_{\mu\nu}\gamma_{AB}$$

$$(\phi_\mu \text{ on atom A}, \phi_\nu \text{ on atom B}) \quad (10b)$$

charge for that atom. γ_{AB} is a spherically averaged interaction between an electron on A and one on B. In the expression for $F_{\mu\nu}$, β°_{AB} is a semiempirical bonding parameter which was chosen initially by calibration with *ab initio* calculations on diatomics.

The molecular orbital equations at the CNDO level are no longer linear, and they have to be solved by iterative methods. Application is consequently more time consuming, but the method is still sufficiently simple for application to quite large molecules with up to ~100 minimal basis functions.

It is clear from eq 10 that the CNDO method takes some account of the influence of charge distribution on the electron-attracting power of an atomic orbital. The second term gives the effect of other electrons on the same atom and the third term allows for stabilization or destabilization due to net charges on neighboring atoms. However, it should be emphasized that other important effects are neglected or poorly taken into account. For example, the method does not include interactions between dipoles due to directed lone pairs of electrons (since atomic distributions are treated as spherically averaged). It is also found that direct bonding energies are considerably overestimated (unless the β_{AB} parameters are substantially reduced).

One easily removed deficiency of the CNDO method is the neglect of one-center exchange integrals. These integrals measure the additional intraatomic stabilization due to electrons with parallel spins occupying different atomic orbitals. This energy contribution leads to additional stabilization for atomic states of high multiplicity (Hund's rule) and is also important in understanding the electron spin distribution in many organic radicals. If contributions from these integrals are added to expressions 10, the semiempirical scheme is described as intermediate neglect of differential overlap (INDO).[8] For most closed-shell diamagnetic molecules, the CNDO and INDO methods give very similar results. The INDO method, however, gives a much superior description of the electron spin density in radicals.

Ab Initio Methods with Slater-Type Orbitals

As indicated in the introduction, we shall consider *ab initio* molecular orbital treatments with two types of minimal basis sets. The first uses functions closely related to Slater-type (exponential) orbitals (STO, eq 11)[9] and similar forms for the other 2p orbitals. The ζ parameters are scale factors which determine the

$$\phi_{1s} = (\zeta_1^3/\pi)^{1/2} \exp(-\zeta_1 r)$$
$$\phi_{2s} = (\zeta_2^5/3\pi)^{1/2} r \exp(-\zeta_2 r)$$
$$\phi_{2p_x} = (\zeta_2^5/\pi)^{1/2} x \exp(-\zeta_2 r) \quad (11)$$

size of the orbital. They may either be chosen as a standard set or may be optimized by the variational method.

Although computational methods are available for the two-electron integrals (eq 8) with Slater orbitals, such calculations are time consuming, so we have replaced the Slater-type functions with linear combinations of Gaussian functions which are chosen to give a least-squares fit. This follows work initiated by Foster and Boys[10] and others.[11,12] Thus we make the replacements

$$\exp(-r) \longrightarrow \sum_{k=1}^{N} a_{1k} \exp(-\alpha_{1k} r^2)$$

$$r \exp(-r) \longrightarrow \sum_{k=1}^{N} a_{2sk} \exp(-\alpha_{2k} r^2)$$

$$x \exp(-r) \longrightarrow \sum_{k=1}^{N} a_{2pk} x \exp(-\alpha_{2k} r^2) \quad (12)$$

where the a and α coefficients are found by methods due to Stewart.[13,14] It should be noted that the expansions of the 2s and 2p functions share common α values. This leads to a substantial gain in computational efficiency. The resulting set of basis functions, with N Gaussians per STO, is referred to as STO-NG. Evaluation of integrals involving sums of Gaussian functions is rapid,[15] and the basis set can be used in quite large molecules.

To test how well the STO-NG basis reproduces the results of a full STO treatment, we may consider the typical total energies listed in Table II. The energies

Table II
Total Energies (Hartrees) with Slater-Type Bases

Molecule	STO-3G	STO-4G	STO
CH_4	−39.715	−39.999	−40.114
C_2H_6	−78.280	−78.841	−79.069
NH_3	−55.454	−55.849	−56.005
H_2CO	−112.330	−113.136	−113.450

for STO-3G and STO-4G are seen to approach the STO values, but only slowly. However, if we consider *atomization energies* (total energy minus atom energies using the same basis set) as listed in Table III, the convergence is seen to be very much more effective. Indeed, atomization energies with the STO-4G basis are almost identical with the STO numbers. This improvement is probably due to the fact that the STO-NG

(8) J. A. Pople, D. L. Beveridge, and P. A. Dobosh, *J. Chem. Phys.*, **47**, 2026 (1967).
(9) J. C. Slater, *Phys. Rev.*, **36**, 57 (1930).
(10) J. M. Foster and S. F. Boys, *Rev. Mod. Phys.*, **32**, 303 (1960).
(11) C. M. Reeves and R. Fletcher, *J. Chem. Phys.*, **42**, 4073 (1965).
(12) K. O-Ohata, H. Taketa, and S. Huzinaga, *J. Phys. Soc. Jap.*, **21**, 2306 (1966).
(13) R. F. Stewart, *J. Chem. Phys.*, **50**, 2485 (1969).
(14) W. J. Hehre, R. F. Stewart, and J. A. Pople, *ibid.*, **51**, 2657 (1969).
(15) S. F. Boys, *Proc. Roy. Soc., Ser. A*, **200**, 542 (1950).

Table III
Atomization Energies (Hartrees) with Slater-Type Bases

Molecule	STO-3G	STO-4G	STO
CH_4	0.586	0.576	0.576
C_2H_6	0.972	0.953	0.952
NH_3	0.309	0.300	0.300
H_2CO	0.349	0.338	0.338

functions deviate most from STO functions near the nucleus, so that the error introduced is similar in a molecule and its separated atoms. Figure 1 shows the STO-3G function for hydrogen which is very close to the exponential form in the outer region where there is overlap with orbitals on neighboring atoms. In most subsequent work we have used the STO-3G basis to permit widespread applications.

The energies listed in Tables II and III mostly used ζ values from Slater's rules for free atoms. To test how close these values are to the best molecular values, complete optimization studies of valence ζ were undertaken for a set of small organic molecules. Some typical results, which parallel other work by Pitzer[16] and by Switkes, Stevens, and Lipscomb,[17] are shown in Table IV. These results indicate that, for hydrogen and

Table IV
Optimum STO Exponents

Molecule	H	C	O
Free atoms	1.00	1.59	2.23
CH_4	1.16	1.76	
C_2H_6	1.17	1.76	
C_2H_4	1.21	1.70	
C_2H_2	1.31	1.67	
H_2O	1.26		2.23

carbon at least, the optimum Slater-type atomic orbitals in molecules are significantly smaller than in free atoms. This was known for hydrogen in the hydrogen molecule in the 1930's, but it is now clear that corresponding effects occur with some other atoms.

Since it is frequently impractical to vary all ζ values for larger molecules, it is useful to define a *standard molecular* set for valence orbitals. Such a set has been proposed and is listed in Table V together with the

Table V
Standard Molecular STO Exponents

Atom	Standard exponent	Slater atom value
H	1.24	1.00
C	1.72	1.625
N	1.95	1.95
O	2.25	2.275
F	2.55	2.6

Slater values for free atoms. In applications of the STO-3G basis, we have usually carried out a fairly full ζ optimization in studies of charge density (which are
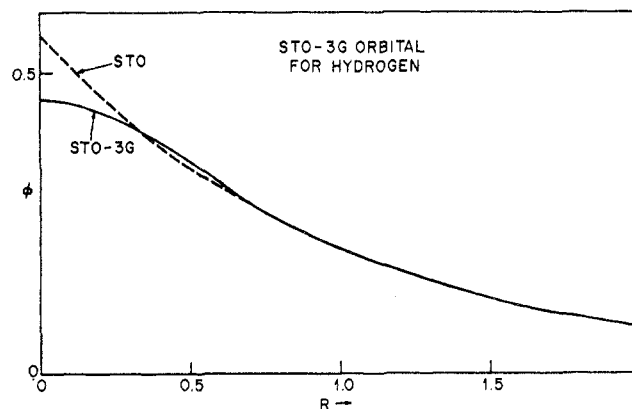

Figure 1. STO-3G orbital for hydrogen.

sensitive to this feature), but for studies of equilibrium geometry only standard ζ values are used.

Ab Initio Methods with Least-Energy Atomic Orbitals

It is well known that, for nonhydrogenic atoms, Slater-type exponential functions are not particularly good approximations to the best (Hartree–Fock) atomic orbitals. It is therefore useful to make a molecular orbital study using minimial basis sets which approach the Hartree–Fock limit for free atoms. We have, therefore, considered another set of atomic orbitals expanded as sums of normalized Gaussian orbitals (eq 13). The d and α parameters could again be chosen by

$$\phi_{1s}' = \sum_{k=1}^{N} d_{1k} \exp(-\alpha_{1k}r^2)$$

$$\phi_2' = \sum_{k=1}^{N} d_{2sk} \exp(-\alpha_{2k}r^2)$$

$$\phi_{2px}' = \sum_{k=1}^{N} d_{2pk}x \exp(-\alpha_{2k}r^2) \quad (13)$$

least-squares fitting, this time with Hartree–Fock orbitals. However, this is found to lead to slow convergence, and a more satisfactory procedure is to vary the d and α until the calculated energy of the atom is minimized. Such a set of functions may then be described as least-energy minimal atomic orbitals with N Gaussians (LEMAO-NG). Such a set of atomic orbitals has recently been determined for $N = 3$–6 and can be used in molecular calculations in a manner similar to the STO-NG sets.[18] By virtue of the way in which they are derived, the atomic energies for LEMAO-NG functions are lower than for STO-NG and converge to the Hartree–Fock limit. This is illustrated in Figure 2.

Just as the STO orbitals change their optimum size in a molecular environment, the least-energy orbitals may also be rescaled, so the actual functions used are given by eq 14, where ζ is now a scaling factor near unity. A number of molecules have been studied with these bases, all ζ-scale factors being optimized.[19]

(16) R. M. Pitzer, *J. Chem. Phys.*, **46**, 4871 (1967).
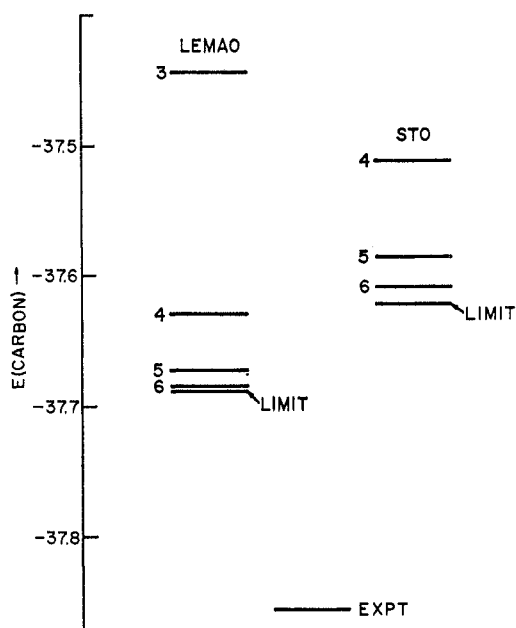(17) E. Switkes, R. M. Stevens, and W. N. Lipscomb, *ibid.*, **51**, 5229 (1969).

(18) R. Ditchfield, W. J. Hehre, and J. A. Pople, submitted for publication.
(19) W. J. Hehre, R. Ditchfield, and J. A. Pople, submitted for publication.

Figure 2. STO and LEMAO atomic energies for carbon.

$$\phi_\mu(r) = \zeta^{3/2}\phi_\mu'(\zeta r) \tag{14}$$

Some typical atomization energies by this method are listed in Table VI, which parallels Table III. The limiting values cannot be obtained, but comparison with LEMAO-6G indicates that convergence of theoretical atomization energies with increasing number of Gaussian functions is less rapid than for the corresponding Slater-type expansions. It is generally necessary to use the LEMAO-4G level to achieve reasonable convergence toward the limiting results.

Table VI
Atomization Energies (Hartrees) with Least-Energy Bases

Molecule	LEMAO-3G	LEMAO-4G	LEMAO-6G
CH_4	0.504	0.495	0.491
C_2H_6	0.843	0.814	0.803
NH_3	0.261	0.259	0.258
H_2CO	0.302	0.277	0.269

Some optimum valence scale factors are shown in Table VII. The most notable feature about these is that the hydrogen functions for all of these molecules are contracted more than in the STO calculations. A rather wider range of scale factors is encountered, and the use of a standard set is rather less satisfactory than for STO. Nevertheless, such a choice can be made and is summarized in Table VIII.

Application to Electron Distribution

All of these molecular orbital methods can be used to study electron distribution in terms of the Mulliken gross populations and associated net atomic charges. It is found that the semiempirical INDO method and the *ab initio* STO-3G method (with optimization of valence ζ values) give fairly comparable distributions. Some typical results for STO-3G are given in Figures 3–8.[20]

(20) W. J. Hehre and J. A. Pople, *J. Amer. Chem. Soc.*, **92**, 2191 (1970).

Table VII
Scaling Factors for Least-Energy 6G Bases

Molecule	H	C	O
CH_4	1.33	1.03	
C_2H_6	1.30	1.05	
C_2H_4	1.33	1.05	
C_2H_2	1.41	1.06	
H_2O	1.51		0.98

Table VIII
Standard Scale Factors for Least-Energy Bases

Atom	Scale factor
H	1.35
C	1.07
N	1.00
O	0.98
F	0.98

These may be used to comment on certain general types of charge displacement that are often discussed in qualitative terms.

Paraffin molecules such as ethane have atomic charges close to zero, corresponding to a nearly nonpolar bond. Ethylene and acetylene show increasingly positive hydrogens corresponding to more $C^- - H^+$ character. The molecule propyne (Figures 4 and 5) is of some interest as it is the simplest hydrocarbon involving the interaction of a methyl group and an unsaturated system. Replacement of an acetylenic hydrogen by methyl leads to a substantial redistribution of charge within the triple bond. It is clear from Figure 4 that this is mainly a polarization leading to an increase of negative character of the β carbon rather than a transfer of charge from the methyl group. Figure 5 gives the corresponding gross populations of the π-type atomic orbitals and indicates that the polarization is mainly associated with π electrons. Figure 6 shows the STO-3G π densities in the phenyl ring of toluene and again show alternating polarization with only a small charge transfer. Thus the STO-3G densities support the qualitative concept of "hyperconjugation" involving π orbitals, but the interaction and consequent polarity are associated with reorganization within the unsaturated system rather than with transfer from the alkyl group.

Figure 7 and Table IX illustrate certain features of

Table IX
Dipole Moments[20]

Molecule	STO-3G	Exptl
Water	1.78	1.85
Methanol	1.51	1.69
Methyl ether	1.18	1.30
Ammonia	1.66	1.47
Methylamine	1.41	1.33
Dimethylamine	1.19	1.03
Trimethylamine	0.95	0.61

the charge distribution associated with the interaction of methyl- and ether-type oxygen. It is apparent that, when attached to oxygen, a methyl group *with-*
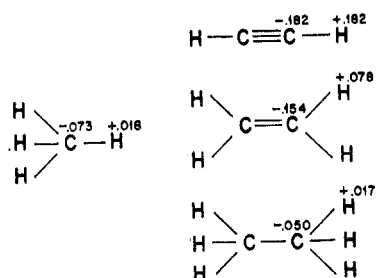
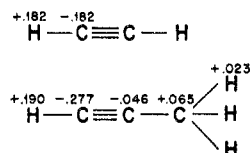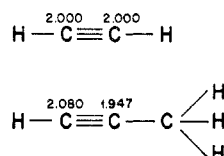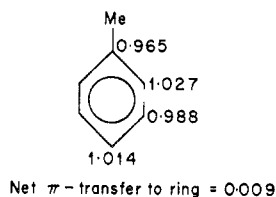Figure 3. STO-3G atom populations for simple hydrocarbons.

Figure 4. STO-3G atom populations for acetylene and propyne.

Figure 5. STO-3G π populations for acetylene and propyne.

Figure 6. STO-3G π populations for toluene.

Figure 7. STO-3G atom populations for water and methyl ether.

Figure 8. STO-3G atom populations for ammonia and tri-methylamine.

Figure 9. Electron transfer in methyl ether.

draws electrons relative to hydrogen according to this treatment. This leads to a dipole moment in methyl ether which is substantially lower than water's, a trend which is fairly well reproduced by the theory. According to the orbital breakdown of the charge distribution, this effect is partly due to back-donation of charge from the π-type lone pair of oxygen into the antibonding orbitals of the methyl group. This electron density shows up on the hydrogens not lying in the oxygen valence plane which become negative in methanol and methyl ether (Figure 9).

Similar effects are found in amines where the decreasing dipole moments of the series NH_3, CH_3NH_2, $(CH_3)_2$-NH, and $(CH_3)_3N$ is associated with decreasing electron density on nitrogen and increasing electron density on the hydrogens *trans* to the nitrogen lone pair (Figure 8 and Table IX). It is interesting to note that the methylamines are stronger bases than ammonia. Thus it appears that, at least for this series, increasing base strength is *not* associated with increasing electron density on nitrogen.

The gross populations using the least-energy minimal atomic orbital set differ markedly from the STO values. The results for methane, ammonia, and water (Figure
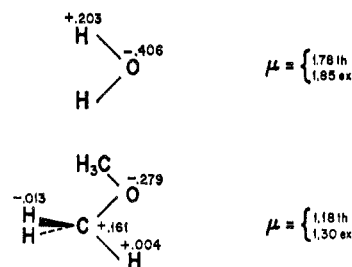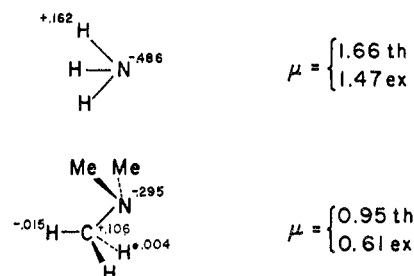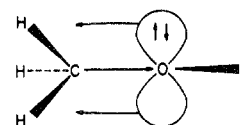
10) show much smaller electron densities on hydrogen and corresponding dipole moments which are considerably too large. It should be noted that this basis set uses hydrogen functions which are considerably more contracted (Table VII), and much of the density in the hydrogen region is probably assigned to the more diffuse functions on the heavy atom. This suggests that a population analysis using a LEMAO basis is less realistic than with the other methods, even though the calculated total energies are lower.

Equilibrium Geometrics

If the standard scale factors listed in Tables V and VIII are used, it becomes possible to carry out extensive geometrical searches for the lowest calculated energy and hence to obtain theoretical equilibrium geometries. A study of this sort using the INDO method has already been published,[21] and the results may now be compared with corresponding studies by the two *ab initio* techniques discussed in this paper.

Table X presents calculated equilibrium geometries for a series of small polyatomic molecules using the INDO, STO-3G, STO-4G, LEMAO-4G, and LEMAO-5G procedures. These results are based on a variation of geometrical parameters which involve a number of constraints such as the requirement that all bonds XH between a heavy atom X and hydrogens H have the same length. However, nonplanar configurations for

(21) J. A. Pople and M. S. Gordon, *J. Chem. Phys.*, **49**, 4643 (1968).

Table X
Equilibrium Geometries[a]

| Molecule	Coordinates	INDO	STO-3G	STO-4G	LEMAO-4G	LEMAO-5G	Exptl
H_2O	r_{OH}	1.03	0.990	0.987	0.986	0.987	0.957
	θ_{HOH}	104.7	100.0	100.0	108.6	109.1	104.5
NH_3	r_{NH}	1.07	1.033	1.029	1.010	1.010	1.012
	θ_{HNH}	106.4	104.2	104.4	114.6	115.0	106.7
CH_4	r_{CH}	1.12	1.083	1.079	1.068	1.068	1.085
C_2H_2	r_{CH}	1.10	1.065	1.063	1.059	1.060	1.061
	r_{CC}	1.20	1.168	1.168	1.197	1.197	1.203
C_2H_4	r_{CH}	1.11	1.079	1.079	1.067	1.066	1.076
	r_{CC}	1.31	1.305	1.307	1.339	1.341	1.330
	θ_{HCH}	111.4	115.4	115.7	116.3	116.3	116.6
C_2H_6	r_{CH}	1.12	1.085	1.083	1.068	1.068	1.096
	r_{CC}	1.46	1.538	1.535	1.577	1.581	1.531
	θ_{HCH}	106.6	108.2	108.2	108.5	108.5	107.8
CH_3F	r_{CH}	1.12	1.097	1.094	1.081	1.081	1.105
	r_{CF}	1.35	1.384	1.381	1.489	1.494	1.385
	θ_{HCH}	109.8	108.3	108.4	112.2	112.3	109.9
CH_2O	r_{CH}	1.12	1.101	1.099	1.087	1.087	1.101
	r_{CO}	1.25	1.217	1.217	1.290	1.293	1.203
	θ_{HCH}	115.0	114.5	114.8	119.6	119.9	116.5
HCN	r_{CH}	1.09	1.070	1.068	1.072	1.073	1.063
	r_{CN}	1.18	1.153	1.154	1.188	1.189	1.154

[a] r given in ångströms, θ in degrees.
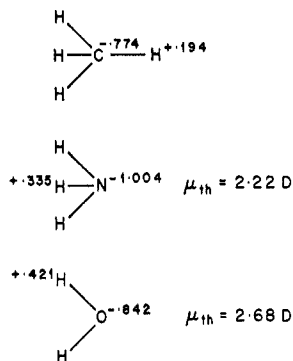
Figure 10. LEMAO-4G atom populations for methane, ammonia, and water.

C_2H_4 and CH_2O and nonlinear configurations for C_2H_2 and HCN were considered, and it does appear very likely that all three methods predict the correct symmetry for all of these molecules.

The first point to note about Table X is that the STO-3G and STO-4G geometries are extremely close and may therefore be considered characteristic of the full STO basis set. The differences between STO-3G and STO-4G geometries are much smaller than the differences between either value and the experimental value, so little is gained by carrying out the calculations at the higher level of Gaussian representation. Similar conclusions follow from a comparison of the LEMAO-4G and LEMAO-5G columns.

To make an overall comparison with experiment, it is useful to examine mean absolute differences between theoretical and experimental geometrical parameters. These are listed in Table XI. From this rather limited sample, it would appear that the semiempirical INDO method is very successful for bond angles but less satisfactory for bond lengths. It is interesting to note that

Table XI
Mean Absolute Errors for Equilibrium Geometries

	INDO	STO-3G	LEMAO-4G
Bond lengths, Å	0.035	0.011	0.028
Bond angles, deg	1.4	2.0	3.1

the Slater-type basis gives better values for both bond lengths and angles than the least-energy atomic orbital set, in spite of the fact that the latter give lower total energies.

A number of detailed points about Table X are worthy of note. Clearly, the prediction of bond angles for water and ammonia is rather sensitive to the method used. The Slater-type basis gives values which are too small and the LEMAO values are too large. Valence angles at carbon, however, are well predicted by all methods. Thus, the fact that the HCH valence angle is less than the trigonal value of 120° in ethylene and formaldehyde is given by all methods, as is the lowering below the tetrahedral value of 109.5° in ethane. The Slater-type basis leads to bond lengths which are in fairly uniform agreement with experiment, but the LEMAO values are too large for C–O and C–F bonds.

On the basis of the results presented in these two tables, it would appear that molecular orbital studies with the simple minimal STO-3G basis and standard exponents should be a valuable technique for predicting the geometrical details of molecular structure. The set of molecules chosen for comparison with experiment is small but fairly representative of organic systems. More extensive investigations are now under way which will include studies of species where the equilibrium geometry is not known experimentally. If the small errors shown in Table XI do hold more generally, theory at this level will become a powerful method of studying molecular geometry.