Theor Chim Acta (1988) 74:275-288

# **Minimal requirements for approximate wavefunctions of molecules\***

9 Springer-Verlag 1988

## **Z. B. Maksi** $\acute{\text{c}}^{1,2}$  **and S. Supek**<sup>2</sup>

<sup>1</sup> Theoretical Chemistry Group, The Rudjer Bošković Institute, Yu-41001 Zagreb <sup>2</sup> Faculty of Natural Sciences and Mathematics, University of Zagreb, Marulićev trg 19, YU-41000 Zagreb, Yugoslavia.

Received May 12, 1987, revised January ll/Accepted January 12, 1988

It is shown that diamagnetic shielding  $\sigma^d$  and diamagnetic susceptibility  $\chi^d$ provide necessary but not sufficient conditions which approximate molecular wavefunctions should satisfy. Hence they represent minimal requirements for acceptable qualitative description of electrons in molecules. Some current semiempirical methods are examined by using  $\sigma^d$  and  $\chi^d$  tests. It appears that the MINDO/3 and MNDO schemes do not perform well for N, O and F atoms. The origin of discrepancies in the calculated  $\sigma^d$  values lies in the use of inadequate orbital screening constants  $\zeta$ . The role of other one-electron properties in improved parametrization of the semiempirical procedures is briefly discussed.

Key words: Diamagnetic shielding — Diamagnetic susceptibility — Independent atom model -- Modified atom model -- Modelling of molecular properties

## **I. Introduction**

Methods currently in use in molecular quantum mechanics are widely different in levels of sophistication and accuracy. The *ab initio* methods with explicit account of correlation effects yield results of high quality for small molecules. However, in large systems one has to resort to approximate and simplified procedures which gain in feasibility by sacrificing some accuracy. Approximate (semiempirical) methods should be as simple as possible and yet they have to reproduce the main features of the electronic structure of molecules. Quantitative

Dedicated to Professor J. Koutecký on the occasion of his 65th birthday

appraisal of molecular wavefunctions calls for suitable criteria. It is well known that total molecular energy, required by the variation theorem to be at minimum, is not a very stringent condition. Namely, the changes in wavefunctions of the first order affect the total energy only in the second order. Substantially better probes are provided by some one- and two-electron properties [1]. They "measure" wavefunctions in different molecular segments. Consequently, testing a particular approximate scheme requires a careful scrutiny of the details of the electronic charge distributions or, in other words, calculations of a number of molecular properties. In this paper we address the question of finding suitable criteria which give insight into qualitative features of approximate wavefunctions. We shall not dwell, of course, on the spatial symmetry and/or permutational requirements for electrons, which certainly have to be built in the acceptable wavefunctions. Instead, an attempt is made to establish criteria which define the lowest threshold in quality of the calculated wavefunctions. If such criteria (properties) exist and the applied methods do not perform in a satisfactory manner, it follows that the underlying assumptions of the approximate schemes are either wrong, inadequate or badly balanced. A closer look will eventually lead to a remedy of the shortcommings and to the improved methods. Since a lower threshold in quality is desired, the corresponding gauge properties should not critically depend on the finer details of the molecular wavefunctions. In particular, they should be rather insensitive to the correlation of electrons. Namely, the semiempirical methods are usually based on the simplified Hartree-Fock hamiltonian and do not involve correlation effects. Hence, if the experimental data are to be used in qualitative examination of the approximate wavefunctions, the desired properties should be one-electronic in nature. The latter do not strongly depend on the correlation [2], the dipole moment being a notable exception [3]. Anticipating results of the present study, one can say that the diamagnetic shielding of the nuclei and the Langevin diamagnetic susceptibility are properties par excellence in determining minimal criteria for acceptable molecular wavefunctions. They provide necessary but not sufficient conditions which good atomic and molecular wavefunctions should satisfy.

#### **2. Selection of qualitative tests for molecular wavefunctions**

Perusal of the *ab initio* results shows that the diamagnetic shielding of the nuclei (Lamb term)  $\sigma^d$  and Langevin's diamagnetic molecular susceptibility  $\chi^d$  are practically unaltered by inclusion of correlation effects [4-6]. Furthermore, they are not very sensitive to details of calculations at the SCF level either, because widely different basis sets give virtually the same  $\sigma^d$  and  $\chi^d$  values. They can be also quite well reproduced by simple formulas based on the independent atom model (IAM). Let us consider first Lamb's term  $\sigma^d$ . It was shown by Ramsey [7] that the average diamagnetic shielding

$$
\sigma_{\text{av}}^d(A) = (e^2/3mc^2)\langle 0|1/r_A|0\rangle\tag{1}
$$

can be expressed in an extremely simple form:

$$
\sigma_{\text{av}}^d(A) = \sigma_{\text{av}}^d(FA) + (e^2/3mc^2) \sum_B' Z_B/R_{AB}
$$
 (2)

where  $\sigma_{av}^{d}(FA)$  denotes the free atom value of atom A and the prime signifies that  $B \neq A$ . The remaining symbols are the well known fundamental constants. It is obvious that the formula (2) corresponds to the IAM or promolecule model where the neutral and spherically symmetric atoms are centered at the equilibrium positions. Ramsey's formula yields reasonable estimates of the diamagentic shielding [8-12] despite the oversimplified physical picture. The free-atom values  $\sigma_{av}^{d}(FA)$  are well described by the atomic Hartree-Fock results. If the intramolecular charge migration is allowed for, the formula (2) takes a form:

$$
\sigma_{\rm av}^d(A) = (e^2/3mc^2) \left[ \sum_{\mu}^A \left( \xi_{A\mu} Q_{\mu}^A / n_{A\mu} \right) + \sum_{B}^{\prime} Q_B / R_{AB} \right]
$$
(3)

where  $Q_{\mu}^{A}$  and  $Q_{A}$  are gross orbital and gross atomic electron populations, respectively. The main quantum number of the corresponding AO is denoted by  $n_{A\mu}$  and  $\xi_{A\mu}$  is the screening constant. In development of the formula (3) it was tacitly assumed that the monopole approximation for the calculation of the  $1/r$ expectation values holds to a good accuracy. A careful analysis of the DZ *ab initio* and semiempirical results has conclusively shown that this supposition is justified [13]. Finally, it should be pointed out that the inner-shell electrons were treated as highly localized and nonpolarizable cores possessing maximal electron occupancy permitted by the Pauli principle. Since the formula (3) involves charge drift between atoms of different electronegativities the underlying physical model is that of modified atoms in molecules (MAM) [12, 14].

Analogous formulas hold for second moments of the electronic change distributions. In the MO-LCAO approach the second moment can be decomposed into three contributions:

$$
\langle r_{\alpha}^{2} \rangle = \sum_{A} \sum_{\mu}^{A} P_{\mu\mu} \langle \phi_{\mu} | r_{\alpha}^{2} | \phi_{\mu} \rangle + 2 \sum_{\mu < \nu}^{A} \sum_{\nu}^{A} P_{\mu\nu} \langle \phi_{\mu} | r_{\alpha}^{2} | \phi_{\nu} \rangle
$$
  
+2 
$$
\sum_{A} \sum_{B} \sum_{\mu}^{B} P_{\mu\nu} \langle \phi_{\mu} | r_{\alpha}^{2} | \phi_{\nu} \rangle
$$
 (4)

where  $\alpha = x, y, z$ , and  $r_{\alpha}$  is the corresponding component of the position vector r. Here  $P_{\mu\nu}$  are elements of the charge density matrix and AOs are denoted by  $\phi_{\mu}$ . A series of consecutive coordinate translations  $r_{\alpha} = r_{A\alpha}(i)$  transforms the formula (4) into a sum of atomic contributions:

$$
\langle r_{\alpha}^{2} \rangle = \sum_{A} \left[ Q_{A} r_{A\alpha}^{2} + \sum_{\mu}^{A} Q_{\mu} \langle \phi_{\mu}(i) | r_{A\alpha}^{2}(i) | \phi_{\mu}(i) \rangle \right] \tag{5}
$$

where  $r_{A\alpha}$  and  $r_{A\alpha}(i)$  signify the  $\alpha$ th coordinate of nucleus A from the origin of the laboratory coordinate system (center of mass) and the corresponding coordinate of the ith electron measured relative to the host nucleus A, respectively. It was assumed in going from (4) to (5) that the Mulliken approximation for mixed densities  $\phi_{\mu} \phi_{\nu} = (1/2)S_{\mu\nu} (\phi_{\mu}^2 + \phi_{\nu}^2)$  holds to a good accuracy. Careful examination of the second term in a large number of molecules has shown that is roughly constant for atoms belonging to the same period of the Mendeleev

system of elements. Thus, the approximate formula (5) takes a lapidary form:

$$
\langle r_{\alpha}^{2} \rangle = \sum_{A} Q_{A} r_{A\alpha}^{2} + \sum_{p} n_{p} k_{p}
$$
 (6)

where  $n_p$  is the number of atoms of the pth row in a molecule and  $k_p$  is a characteristic constant. It is interesting that the empirical  $k_p$  values correspond very closely to the *ab initio* data for free atoms [15] averaged over the pth period of the Mendeleev system. Hence the second moments can be expressed in a very simple way in a parameter-free form. Needless to say the formula (6) corresponds to the MAM picture of molecules. The approximate constancy of the second term in (5) can be rationalized by an observation that the more electronegative atoms have larger populations  $Q<sub>u</sub>$  but at the same time the corresponding integrals  $\langle \phi_{\mu} | r_{\alpha}^2 | \phi_{\mu} \rangle$  are smaller. The latter measure the size of the atomic orbitals which is smaller in atoms with higher  $Z_A$  within the same period which is a consequence of the higher nuclear-electron attraction. By neglecting intramolecular charge distribution and putting  $Q_A = Z_A$  one obtains:

$$
\langle r_{\alpha}^{2} \rangle = \sum_{A} Z_{A} r_{A \alpha}^{2} + \sum_{p} n_{p} k_{p}
$$
 (7)

This is the IAM form of the approximate formula (6). It should be noticed that the second moment is intrinsically a negative entity. The negative sign is, however, dropped for simplicity.

Extensive calculations of Lamb's shielding and Langevin's susceptibility have shown that simple and transparent formulas  $(2, 3)$  and  $(6, 7)$  are adequate for most purposes. Interestingly, the IAM model has a surprisingly good performance [8-14]. Hence, the  $\sigma^d$  and  $\chi^d$  properties can be well reproduced by using models of utmost simplicity. Consequently, it follows that they provide minimal conditions which approximate wavefunctions should satisfy. The approximate but appropriate wavefunctions should yield  $\sigma^d$  and  $\chi^d$  values which are close to experimental data or good *ab initio* results. They are not satisfactory otherwise.

## **3. Examination of some current semiempirical methods by using**  $\sigma^d$  **and**  $\chi^d$  **tests**

The customary all-valence semiempirical methods can be divided into two categories: (a) schemes which are based on some variants of the NDDO approximation and (b) methods which fully appreciate overlapping of atomic orbitals. Typical representatives of the first class are MINDO/3 [16] and MNDO [17] methods. The EHT [18], SCC-MO (or IEHT) [19] and iterative maximum overlap (IMO) [20] methods belong to the second class of approximate procedures. It would be of some interest to explore behaviour of the MINDO/3 and MNDO methods in reproducing  $\sigma^d$  and  $\chi^d$  values. Namely, various NDDO schemes could be justified to a considerable extent if proper bases of suitably orthogonalized hybrid orbitals are employed [21-23]. Since this is not the case, the MINDO/3 and MNDO schemes are inconsistent. This fact should be reflected in the calculated  $\sigma^d$  and  $\chi^d$  values. As a method which is based on the model hamiltonian and explicitly takes into account orbital overlapping we shall choose the SCC-MO approach in a form prescribed by Van der Voorn and Drago [19]. As a basis set

functions we shall use Clementi-Raimondi orbitals [24]. They appreciate a difference in screening of s and p types of AOs. The MINDO/3 and MNDO methods were utilized in their original versions. These two methods are parametrized to reproduce not only heats of molecular formation but also structural data of molecules. It is often argued that their wavefunctions are better if they refer to optimized geometries instead of experimental ones. We have executed MINDO/3 and MNDO calculations for both optimized and observed structural parameters. Since the validity of formulas (3) and (6) is fully established, they will be employed in the present calculations. In addition to computer-time saving they offer a simple interpretation of the studied properties. We note in passing that the use of approximate formulas for one-electron property matrix elements is not always justified. For example, semiempirical estimates of electric field gradients at the nuclei require rigorous calculation of the corresponding integrals [25]. The *ab initio* results of Snyder and Basch [26] are selected as gauge values for  $\sigma^d$  and  $\chi^d$  physical properties. They are of the DZ quality thus being relatively close to true values. Their changes caused by enlargement of the basis set is very small [26]. The use of Snyder and Basch calculations ensures uniformity of results for a large set of molecules. Finally, the IAM estimats of  $\sigma^d$  and  $\chi^d$  properties are given for the sake of comparison and as an illustration of their insensitivity.

Diamagnetic shieldings of the nuclei in 41 widely different molecules are calculated involving first row atoms in various chemical environments. They are summarized in Tables 1 and 2. In addition to IAM, MINDO/3, MNDO and SCC-MO results, diamagnetic shieldings of Ray and Parr [27] are included. They were obtained via an exact expression involving a nuclear term, a derivative of the total energy over the nuclear charge and the chemical potential (i.e. the negative of the molecular electronegativity). The total energy was estimated by a simple empirical two-parameter formula of Gadre and Parr [28].

$$
E_t = -0.51130 \ (Z_A^{2.40073} + Z_B^{2.40073})
$$

generalized in an obvious way for polyatomic molecules. We note in passing that the formula of Gadre and Parr falls within the modified atom in molecule (MAM) model. Two types of results are arrayed in the last but one columns in Tables 1 and 2. The first refers to diamagnetic shieldings offered by the neglect of the chemical potential. They are less satisfactory as we shall see soon. The second, given within parentheses, is obtained by inclusion of the chemical potential contribution. Inspection of the results indicates that the IAM model is the closest to *ab initio* values. The SCC-MO procedure comes next. The corresponding average absolute error  $\Delta_a$  are 1.5 and 1.9 (in ppm), respectively. The Ray-Parr formula with ommission of the chemical potential yields a relatively high average absolute error ( $\Delta_a = 5.8$  ppm). Full appreciation of the chemical potential leads to a decrease of the average absolute error by 57% ( $\Delta_a = 2.5$  ppm). Apparently, chemical potential is important and can not be abandoned. MINDO/3 and MNDO procedures obviously do not perform very well as evidenced by the  $\Delta_a$ values. They are 10.0 and 9.1 ppm, respectively. The MNDO is somewhat better than the MINDO/3 method as expected, because its semiempirical scheme is more versatile. The largest discrepancy of MINDO/3 approach is found for

Molecule	Gauge origin	$MINDO/3^a$	MNDO <sup>a</sup>	SCC-MO	<b>IAM</b>	Ray-Parr <sup>b</sup>	Ab initio $DZ^c$
H <sub>2</sub>	Н	$35.8$ $(35.6)$	$36.3$ $(37.8)$	33.9	30.4	34.5 (29.9)	32.2
$N_{2}$	N	392.9 (392.7)	396.7 (396.1)	383.0	385.4	391.3 (387.9)	384.1
F <sub>2</sub>	F	484.7 (483.3)	543.7 (550.6)	525.4	530.7	533.4 (526.6)	529.9
HF	H	102.7 (110.9)	104.1(112.0)	102.9	110.2	114.1(108.5)	107.9
	$\mathbf{F}$	435.5 (436.0)	497.8 (498.6)	478.3	481.3	483.9 (478.3)	482.2
$_{\rm CO}$	C	327.3 (326.7)	330.2 (327.6)	324.6	327.3	335.1 (330.2)	326.1
	O	463.9 (463.6)	468.2 (467.3)	444.9	445.0	451.6 (446.8)	445.1
BF	B	262.8 (266.4)	272.3 (271.9)	265.4	269.0	274.8 (270.5)	267.4
	F	463.1 (464.7)	523.8 (523.7)	507.0	508.2	510.8 (506.5)	507.8
$C_2H_2$	H	102.7(101.7)	102.6(103.0)	101.8	98.6	$102.8$ (98.6)	99.1
	C	323.5 (323.4)	326.4 (326.6)	319.2	320.6	328.4 (324.2)	321.5
<b>HCN</b>	н	105.8(102.8)	103.2(103.3)	102.4	100.3	104.6(100.1)	99.6
	$\mathbf C$	327.5 (327.5)	331.2 (330.8)	324.2	326.4	334.3 (329.7)	326.6
	N	385.7 (386.4)	391.7 (391.3)	377.8	378.5	386.2 (381.7)	378.6
CO <sub>2</sub>	C	357.4 (382.6)	391.0 (387.7)	386.0	390.1	398.0 (392.7)	386.8
	O	495.8 (494.4)	499.3 (497.3)	475.4	476.0	482.6 (477.3)	476.2
N(1)N(2)O	N(1)	424.4 (425.7)	427.2 (427.0)	412.9	416.3	424.0 (418.5)	414.4
	N(2)	444.3 (444.8)	454.6 (454.3)	442.7	447.2	455.0 (449.5)	443.1
	O	501.2(501.0)	503.8 (508.6)	479.8	479.1	485.7 (480.2)	479.6
C(1)(C(2)O) <sub>2</sub>	C(1)	415.2 (413.9)	419.5 (418.9)	404.9	411.0		410.3
	C(2)	408.4 (407.1)	412.8 (411.5)	407.0	412.4	$\overline{\phantom{0}}$	408.9
	O	517.1 (516.4)	519.7 (518.6)	496.1	497.5		497.1

Table 1. Comparison of the diamagnetic shielding of the nuclei in some diatomic and linear molecules as computed by several semiempirical methods, independent atom model (IAM) and *ab initio* DZ approach (in ppm)

Entries in parentheses refer to optimized geometries. Experimental structural data given in [26] are used otherwise

<sup>b</sup> Taken from [27]. Entries in parentheses were obtained by taking into account electronegativity correction

 $\degree$  [26]

fluorine atom. This is not surprising because this atom is not well parametrized  $[16]$ . The next atom which is not satisfactorily described in the MINDO/3 scheme is oxygen. The MNDO procedure does not perform well for fluorine, nitrogen and oxygen, the latter atom being the worst case. It is comforting that most of the errors are systematic in nature. This is particularly true for the MNDO procedure. The average absolute errors for N, O and F atoms of the MINDO/3 (MNDO) methods are: 6.6 (12.8), 15.5 (22.9) and 45.3 (15.0), respectively (in ppm). Survey of the data shows that clustering of the errors around the average value is very dense for the MNDO method. The MINDO/3 results are much more scattered being sometimes erratic. For example, scattering of the MINDO/3 absolute errors in  $\sigma_{av}^{d}(O)$  values ranges from  $\sim$ 7 ppm to  $\sim$ 22 ppm. The same holds for absolute errors of nitrogen shieldings which assume values between  $\sim$ 1 ppm and  $\sim$ 10 ppm. Apparently, the MNDO method describes the studied systems in a more uniform way. The question arises what is the origin of these substantial discrepancies. The formula (3) shows that  $\sigma_{av}^{d}(A)$  values depend on

Molecule	Gauge origin	$MINDO/3^b$	MNDO <sup>b</sup>	SCC-MO	IAM	Ray-Parr <sup>e</sup>	Ab initio DZ
BH <sub>3</sub>	Н	$71.6$ $(71.7)$	72.6 (74.2)	69.9	66.2	70.2 (66.1)	68.6
	B	218.3 (218.4)	228.6 (229.4)	222.3	225.6	231.5 (227.4)	226.4
CH <sub>4</sub>	Н	90.4 (89.8)	91.0 (90.2)	88.6	85.0	89.2 (84.8)	87.1
	$\mathbf C$	296.5 (296.1)	298.6 (298.2)	293.6	295.1	302.8 (298.4)	296.2
H <sub>2</sub> O	Н	105.3(105.9)	106.7 (107.7)	104.7	102.4	106.6(101.5)	102.1
	$\mathbf O$	435.4 (435.5)	439.7 (440.0)	415.9	414.7	421.3 (416.1)	416.8
$B_2H_6$	$H_t$				94.6	98.6 (94.5)	97.0
	H <sub>b</sub>				110.9	114.9 (110.8)	111.2
	B				265.6	271.4 (267.3)	266.0
$C_2H_4$	н	113.4 (112.2)	113.7 (113.1)	111.6	108.4	112.4 (108.0)	110.0
	$\mathbf C$	330.9 (331.1)	333.9 (333.7)	327.5	329.4	336.8 (332.4)	330.1
$N_2H_2$	н	125.4 (123.6)	125.1(125.1)	123.1	120.5	124.8 (119.9)	120.9
	$\mathbf N$	400.2 (401.3)	405.7 (437.4)	391.9	393.1	400.8 (395.9)	392.9
$H_2CO$	H	114.8 (114.4)	115.0(115.1)	112.0	110.1	114.4 (109.6)	112.3
	$\mathsf{C}$	337.0 (338.0)	341.5 (341.0)	336.7	339.7	347.6 (342.8)	338.8
	$\circ$	471.9 (472.1)	476.2 (475.7)	453.1	451.0	457.6 (452.9)	452.7
$C_2H_6$	H	121.2 (115.4)	121.7 (115.9)	119.3	115.7	120.0(115.6)	118.1
	$\mathbf C$	336.9 (338.2)	339.7 (340.0)	334.2	335.7	343.6 (339.1)	337.2
CH <sub>4</sub>	H	90.4 (89.8)	91.0 (90.2)	88.6	85.0	89.2 (84.8)	87.2
	Ċ	296.5 (296.1)	298.6 (298.2)	293.6	295.1	302.8 (298.4)	296.7
$N_2H_4$	H(1)	132.8 (131.4)	132.8 (133.5)	131.1	128.0	132.3 (127.4)	129.0
	H(2)	134.2 (131.3)	134.4 (131.3)	132.3	129.3	133.6 (128.8)	129.9
	N	404.3 (408.9)	410.9 (415.0)	396.2	397.1	404.9 (400.0)	396.5
$H_2O_2$	Н	144.7 (142.0)	146.0 (144.0)	143.7	142.2	146.5 (141.1)	141.4
	о	468.9 (471.5)	484.2 (491.7)	461.3	461.1	467.7 (462.3)	462.0
$CH(1)H(2)_{2}$	H(1)	124.7 (133.5)	125.0 (124.9)	121.6	120.3	124.2 (119.5)	122.1
OH(3)	H(2)	124.8 (132.2)	125.2 (123.7)	122.3	119.4	123.4 (118.7)	121.2
	H(3)	139.9 (145.5)	141.1 (141.5)	139.8	136.5	140.7(136.0)	136.4
	C	343.0 (345.7)	346.6 (347.4)	341.2	344.1	351.7 (347.0)	344.2
	О	466.6 (477.3)	483.6 (484.6)	460.9	457.9	464.6 (459.9)	460.3
CH <sub>3</sub> F	н	125.8 (125.7)	126.1 (124.9)	122.8	121.4	125.4 (120.6)	122.8
	C	345.2 (345.3)	349.2 (350.3)	345.2	347.9	355.3 (350.5)	347.2
	F	480.9 (481.3)	520.8 (544.0)	527.0	525.7	528.3 (523.5)	527.7
$H_2C(1)C(2)O$	H	130.5 (129.2)	131.0 (129.7)	128.1	126.2	130.1(125.6)	126.6
	C(1)	355.1 (354.7)	357.1 (356.3)	346.7	351.8	359.1 (354.6)	352.1
	C(2)	374.6 (373.9)	379.8 (378.0)	373.4	378.0	385.3 (380.7)	375.8
	O	492.2 (491.8)	495.1 (493.9)	472.0	472.5	479.0 (474.5)	473.2
CH <sub>2</sub> N(1)N(2)	Н	133.3(131.6)	133.2(131.8)	131.0	128.5	132.5 (127.9)	129.5
	C	359.2 (360.8)	360.5 (360.2)	351.0	355.3	362.7 (358.0)	355.2
	N(1)	436.3 (437.5)	446.4 (445.5)	432.3	436.8	443.8 (439.2)	433.9
	N(2)	422.0 (422.7)	425.3 (424.0)	411.1	414.1	421.1 (416.5)	412.5
H <sub>2</sub> N(1)CN(2)	H	147.3 (140.3)	146.7 (141.0)	145.0	143.7	147.7 (143.0)	143.0
	C	376.4 (376.5)	380.4 (380.1)	372.4	376.1	383.4 (378.7)	375.2
	N(1)	418.1 (417.5)	427.6 (426.6)	409.7	414.9	421.9 (417.2)	414.1
	N(2)	415.0 (415.3)	419.1 (419.5)	404.8	406.3	413.4 (408.7)	406.3
BH <sub>3</sub> CO	н	121.5(122.3)	123.5 (124.1)	119.5	116.2	120.2 (115.8)	118.6
	B	283.1 (285.7)	296.8 (298.3)	286.7	290.5	296.4 (292.0)	292.1
	C	367.1 (368.5)	369.8 (368.6)	365.9	371.1	378.5 (374.1)	368.2
	$\mathbf O$	477.1 (489.7)	491.3 (490.3)	468.9	471.4	478.0 (473.6)	470.6

Table 2. Comparison of the diamagnetic shielding of the nuclei in some medium size molecules as obtained by several semiempirical methods, independent atom model (IAM) and *ab initio* DZ approach (in  $ppm)$ <sup> $a$ </sup>





 $a$  In azirine cis- and trans-protons are denoted by  $H(1)$  and  $H(2)$ , respectively

b Entries in parentheses refer to optimized geometries. Experimental geometries given in [26] are used otherwise

r Taken from the [27]. Entries given in parentheses were obtained by taking into account electronegativity correction

orbital screening constants  $\zeta$  and orbital populations. The former have a predominating effect as evidenced by the following analysis. Let us consider simple homodiatomic molecules  $N_2$  and  $F_2$ . Interatomic charge transfer is zero by symmetry. The only redistribution of electron population takes place between 2s and 2p orbitals. The screening constants  $\zeta$  are optimized in the MINDO/3 and MNDO methods by reproducing some observables  $(\Delta H_f,$  geometry etc.) together with a number of other parameters [16, 17]. Their numerical values are at considerable variance with other prescriptions, e.g. Clementi-Raimondi [24] and Burns [27] values. We shall, therefore, use Clementi-Raimondi  $\zeta$  constants, which are quite successful within the SCC-MO framework, and retain MINDO/3 and MNDO orbital populations. This simple approach will shed some light on the role of the screening parameters. Improvement of the results is dramatic. In  $N_2$ the MINDO/3 and MNDO  $\sigma_{av}^{d}(N)$  values are both now 380.6 ppm. Further increase in accuracy is obtained by the use of Ransil's best atom and best limited  $\xi$  [28]. The latter distinguish between the p $\sigma$  and  $p\pi$  orbitals which seems to be of importance in linear systems at the least. The MINDO/3 (MNDO) results based on the best atom  $\zeta$  are 381.8 (381.8), whereas the Ransil best limited  $\zeta$ yield 385.2 (385.0)ppm in excellent agreement with *ab initio* DZ value (Table 1). Similar improvement is found in the  $F_2$  case. Ransil's best atom and Clementi-Raimondi basis sets give  $\sim$  525 ppm for  $\sigma_{av}^d(F)$  for both MINDO/3 and MNDO methods. This is reasonably close to the *ab initio* result of 529.2 ppm. Let us focus attention to some heterodiatomics where the charge transfer is highly pronounced. In BF molecule the MINDO/3 (MNDO) methods with Clementi-Raimondi  $\xi$  yield  $\sigma_{av}^d(F)$  = 507.1 (504.8) ppm. The corresponding Ransil's best atom results are 507.4 (505.1) ppm again in fine accordance with *ab initio*  computations. Boron shielding is predicted to be 265.5 (265.9) and 266.6 (266.9) ppm for Clementi-Raimondi and Ransil's best atom  $\xi$ , respectively, the MNDO result being cited always in parentheses. Degree of agreement with GTO DZ results is remarkable indeed. Carbon monoxide is considered as a last example. Clementi-Raimondi and Ransil's best atom  $\xi$  give  $\sigma_{av}^{d}(O)$  values of 443.3 (442.1) and 444.0 (442.8) ppm offering thus a substantial improvement in description of the recalcitrant oxygen atom. Hence, it follows that discrepancies of the MINDO/3 and MNDO results should be ascribed to the use of inadequate  $\zeta$ . This finding is of great importance because  $\xi$  constants strongly influence the  $1/r_A$  operator which plays a crucial role in energetics of atoms and molecules.

Diamagnetic shieldings of other atoms are relatively well reproduced although there is room for some improvement too (Tables 1, 2). Optimized geometries did not change the overall picture offered by the original MINDO/3 and MNDO methods to a significant extent. In some cases additional discrepancies appeared due to inaccurate structural parameters.

A point of considerable interest is observation that  $\sigma^d$  values strongly depend of  $\xi$ 's of the atom in question but not on the orbital populations. This conjecture is substantiated by a comparison of orbital populations in BF:  $Q_{2s}^B = 1.982$  (1.889),  $Q_{2px}^B = 0.107 (0.601), \tQ_{2pz}^B = 0.309 (0.181), \tQ_{2s}^F = 1.941 (1.815), \tQ_{2px}^F =$ 1.971 (1.696),  $Q_{2py}^r = Q_{2pz}^r = 1.692$  (1.819), where the MNDO results are placed in

parentheses as usual. In spite of widely different populations the resulting  $\sigma^d$ **values are very close and are in fact quite acceptable if Clementi-Raimondi or Ransil's**  $\xi$  **are employed. The same holds for CO molecule. Therefore, it appears that diamagnetic shielding provides a** *necessary but not sufficient* **criterion which approximate wavefunctions should satisfy.** 

**The second moments are displayed in Tables 3 and 4. Apart some "pathological" molecules (BF, HCN and HNO) all three semiempirical methods and the IAM model reproduce the DZ results in a satisfactory manner. Their absolute average**  errors are  $0.4 \times 10^{-16}$  cm<sup>2</sup>, if the experimental geometries are used. The errors are **practically doubled if MINDO/3 (0.8) and MNDO (0.7) optimized geometries are employed, because prediction of structural parameters is moderately good. It is noteworthy that the second moments do not depend on the coupling between s and p AOs (hybridization) in the first row atoms in contrast to dipole moments. The same holds for the coupling of s and d AOs in atoms of the second and higher rows of the system of elements. This conjecture is substantiated by the good performance of the formulas (6) and (7) in a wide variety of molecules involving heavier atoms [31, 32]. Further, since the average values of the second moment operator depend on the inverse square of the orbital screening constants,** 

Molecule	IAM	MINDO/3	<b>MNDO</b>	SCC-MO	Ab initio DZ
H <sub>2</sub>	$\langle x^2 \rangle = 0.4$	0.4	0.4	0.4	0.408
	$\langle z^2 \rangle = 0.68$	0.67(0.68)	0.67(0.62)	0.67	0.592
HF	$\langle x^2 \rangle = 1.2$	1.2	$1.2^{\circ}$	1.2	1.152
	$\langle z^2 \rangle = 2.0$	1.70(1.70)	1.76(1.81)	1.79	1.465
$N_{2}$	$\langle x^2 \rangle = 2.0$	2.0	2.0	2.0	2.112
	$\langle z^2 \rangle$ = 6.2	6.20(6.20)	6.20(6.26)	6.20	6.814
$_{\rm CO}$	$\langle x^2 \rangle = 2.0$	2.0	2.0	2.0	2.085
	$\langle z^2 \rangle = 6.4$	6.29(6.35)	6.31(6.60)	6.31	7.064
BF	$\langle x^2 \rangle = 2.0$	2.0	2.0	2.0	2.134
	$\langle z^2 \rangle = 7.1$	5.00(6.50)	5.00(7.08)	5.00	8.320
F <sub>2</sub>	$\langle x^2 \rangle = 2.0$	2.0	2.0	2.0	1.965
	$\langle z^2 \rangle = 11.0$	11.01(11.41)	11.01(9.21)	11.01	10.924
$C_2H_2$	$\langle x^2 \rangle = 2.4$	2.4	2.4	2.4	2.888
	$\langle z^2 \rangle = 12.3$	11.68(11.72)	11.47 (11.38)	12.05	11.227
HCN	$\langle x^2 \rangle = 2.2$	$2.2^{\circ}$	2.2	$2.2^{\circ}$	2.462
	$\langle z^2 \rangle = 8.6$	6.80(9.21)	6.81(8.78)	6.81	8.999
CO <sub>2</sub>	$\langle x^2 \rangle = 3.0$	3.0	3.0	3.0	3.072
	$\langle z^2 \rangle = 24.5$	25.77 (26.60)	25.08 (25.90)	24.99	26.075
<b>NNO</b>	$\langle x^2 \rangle = 3.0$	3.0	3.0	3.0	3.197
		$\langle z^2 \rangle$ = 23.1 21.69 (23.77)	21.08 (23.38)	20.96	24.331
$C_3O_2$	$\langle x^2 \rangle = 5.0$	5.0	5.0	5.0	5.493
		$\langle z^2 \rangle$ = 119.9 121.93 (123.58) 119.79 (121.02)		121.82	122.521

**Table 3. Comparison of second moments of the electronic charge distributions in some diatomics and linear molecules as obtained by several current semiempirical methods, IAM model and** *ab initio*  DZ approach<sup>9</sup> (in  $10^{-16}$  cm<sup>2</sup> units)

**<sup>a</sup>The second moments are computed relative to the center of mass. Values given within parentheses refer to optimized geometries obtained by the method given on the column's heading. Experimental structural data are used otherwise** 

Molecule	IAM	MINDO/3	<b>MNDO</b>	SCC-MO	Ab initio DZ
BH <sub>3</sub>	$\langle x^2 \rangle = 3.7$	3.78 (3.77)	3.89(3.48)	3.83	4.059
	$\langle v^2 \rangle = 1.6$	1.6	1.6	1.6	1.481
	$\langle z^2 \rangle$ = 3.7	3.78 (3.77)	3.89 (3.75)	3.83	4.059
BF <sub>3</sub>	$\langle x^2 \rangle = 26.6$	27.66 (28.03)	26.97 (27.84)	27.11	27.456
	$\langle y^2 \rangle = 4.0$	4.0	4.0	4.0	3.689
	$\langle z^2 \rangle$ = 26.6	27.66 (28.03)	26.97 (27.84)	27.11	27.456
$H_2O$	$\langle x^2 \rangle = 1.4$	1.4	1.4	1.4	1.530
	$\langle y^2 \rangle = 2.5$	2.25(2.23)	$2.36$ $(2.36)$	2.36	2.014
	$\langle z^2 \rangle = 2.0$	1.84 (1.84)	1.89(1.85)	1.89	1.814
${\bf O}_3$	$\langle x^2 \rangle = 22.0$	22.57 (23.73)	22.34 (20.10)	22.06	22.696
	$\langle y^2 \rangle = 3.0$	3.0	3.0	3.0	3,076
	$\langle z^2 \rangle$ = 5.4	5.29 (4.65)	5.32 (4.97)	5.36	5.74
CF <sub>2</sub>	$\langle x^2 \rangle = 22.1$	22.952	22.45 (23.46)	22.40	22.733
	$(y^2) = 3.0$	3.0	3.0	3.0	2.882
	$\langle z^2 \rangle$ = 5.8	5.35	$5.69$ $(3.62)$	5.70	6.442
<b>FNO</b>	$\langle x^2 \rangle = 14.5$	14.7 (4.86)	14.58 (11.08)	14.54	14.933
	$\langle y^2 \rangle = 3.0$	3.0	3.0	3.0	2.946
	$\langle z^2 \rangle$ = 14.9	15.27 (27.80)	15.01 (15.66)	14.94	15,443
$H_2CO$	$\langle x^2 \rangle = 4.2$	4.35 (4.20)	4.19(4.10)	4.08	4.158
	$\langle y^2 \rangle = 2.4$	2.4	2.4	2.4	2.446
	$\langle z^2 \rangle = 10.4$	10.45 (10.72)	10.27 (10.55)	10.14	10.428
$F_2CO$	$\langle x^2 \rangle = 4.0$	4.0	4.0	4.0	3.884
	$\langle y^2 \rangle$ = 25.1	26.54 (19.68)	26.00 (25.79)	26.01	26.362
	$\langle z^2 \rangle = 24.3$	25.24 (30.75)	24.66 (25.91)	24.53	25.257
CHOF	$\langle x^2 \rangle$ = 24.1	24.87 (24.53)	24.43 (24.07)	24.39	25.020
	$\langle y^2 \rangle = 3.2$	3.2	3.2	3.2	3.163
	$\langle z^2 \rangle$ = 7.1	7.29 (8.97)	7.00(25.58)	6.93	6.974
CH <sub>2</sub> CO	$\langle x^2 \rangle$ = 5.2	4.99 (4.92)	5.03(4.94)	5.11	4.978
	$\langle y^2 \rangle = 3.4$	3.4	3.4	3.4	4.038
	$\langle z^2 \rangle$ = 31.2	31.71 (32.31)	31.16 (32.21)	31.13	31.721
CH <sub>4</sub>	$\langle x^2 \rangle = 3.3$	3.40(3.43)	3.40(3.45)	3.32	3.330
$CF_4$	$\langle x^2 \rangle$ = 25.9	26.90 (26.36)	26.39 (27.27)	26.34	26.519
$CH_3F$	$\langle x^2 \rangle = 4.2$	4.30 $(4.36)$	4.20 (4.25)	4.09	4.024
	$\langle y^2 \rangle$ = 4.2	4.30(4.36)	4.20 $(4.25)$	4.09	4.024
	$\langle z^2 \rangle$ = 13.2	13.32 (9.88)	13.20 (12.94)		
CHF <sub>3</sub>	$\langle x^2 \rangle$ = 25.3	26.37 (25.96)	25.94 (26.15)	13.04	13.169
	$\langle y^2 \rangle = 25.3$	26.37 (25.96)		25.96	26.136
	$\langle z^2 \rangle$ = 7.5	7.35 (7.29)	25.94 (26.15)	25.96	26.136
NH <sub>3</sub>	$\langle x^2 \rangle = 2.9$	2.81 (2.85)	$7.07$ $(7.45)$ 2.77	6.90	6.828
	$\langle y^2 \rangle = 2.9$	2.81(2.85)	(2.78) 2.77 (2.78)	2.84	2.599
	$\langle z^2 \rangle = 1.9$	1.90(1.99)		2.84	2.599
$N_2H_4$	$\langle x^2 \rangle$ = 14.5	14.41 (13.82)	1.89(1.94) 14.34 (13.58)	1.91	2.212
	$\langle y^2 \rangle$ = 4.7			14.38	14.789
	$\langle z^2 \rangle$ = 4.7	4.61 (4.91) (4.07)	4.53 (4.78)	4.58	4.526
$N_2F_2$	$\langle x^2 \rangle = 26.8$	4.61 27.41	4.53(4.05)	4.58	4.526
	$\langle y^2 \rangle$ = 36.6	38.34	26.94 (29.46)	27.01	27.553
	$\langle z^2 \rangle = 4.0$		36.91 (31.91)	37.13	37.411
$N_2H_2$	$\langle x^2 \rangle = 9.7$	4.0	4.0	4.0	4.052
	$\langle y^2 \rangle$ = 4.2	9.65 (9.28) 4.10(4.31)	9.56 (9.41) 3.98	9.7	9.978
	$\langle z^2 \rangle = 2.4$	2.4	(4.04) 2.4	4.2	3.977
				2.4	2.536

Table 4. Comparison of the second moments of the electronic charge distributions in some medium size molecules as obtained by several semiempirical methods, IAM model and *ab initio* DZ approach  $(in 10^{-16} cm^2 units)$ 





they are not very sensitive on the latter. This is easily seen by the Taylor expansion  $1/(\xi + \Delta \xi)^2 = (1/\xi^2) - 2\Delta \xi/\xi^3 + \cdots$ . However, it is somewhat surprising that the second moments are quite insensitive to the distribution of monopoles of the electronic charge distributions  $Q_A$  too. In any case, the second moments offer the second necessary but not sufficient condition for approximate wavefunctions. It is interesting to mention that some  $X_\alpha$  calculations do not satisfy this condition and that the second moments can be used as a means in determining a degree of overlapping of the atomic domains [14].

### **4. Final remarks**

We have shown that  $\sigma^d$  and  $\chi^d$  tests provide a rough filter for approximate atomic and molecular wavefunctions. Analysis of the results yields useful information about methods applied. For example, it shows that molecular electronegativity can not be dropped in Ray-Parr formula for diamagnetic shielding [27]. Similarly, imperfections in  $X_{\alpha}$ -method, LCAO- $X_{\alpha}$  procedure and Thomas-Fermi-Dirac method are detected by examining their performance in reproducing HF values of  $\sigma^d$  and  $\chi^d$  properties [14, 33]. Thus, these properties should be used in obtaining better versions of approximate methods. Furthermore, careful examination of the present results reveals that diamagnetic shielding  $\sigma^d$  is strongly dependent on the orbital screening exponents and to a much less extent on the orbital populations. Hence,  $\sigma^d$  can serve as a useful aid in determining basis sets possessing appropriate screening parameters. This is of some importance because some of the most successful semiempirical methods like MINDO/3 and MNDO obviously employ unrealistic  $\xi$  values at least for N, O and F atoms. The latter were obtained by optimizing the calculated heats of formation and structural parameters by treating all adjustable constants on an equal footing. In fact, AOs' screening parameters should be separately determined by reproducing experimental or good *ab initio*  $\sigma^d$  shieldings in a set of characteristic gauge molecules. An approach like this would enable a controlled parametrization with physical justification of parameters. Such a systematic and perhaps stepwise parametrization of semiempirical schemes would be advantageous because simultaneously optimized adjustable parameters are interlocked and influence each other in an unpredictable way particularly if their number is very large. It is clear that other molecular properties are necessary for that purpose. They include not only the most important observables as molecular heats of formation and structural data, but involve in addition physical properties like ESCA chemical shifts and electric field gradients. The latter give useful information about the population of the p-subshell of the atom in question  $[25]$ , whilst the former provide the most direct insight into the distribution of atomic monopoles [34]. Knowledge of the accuracy of one- and two-electron components of the total energy might ve valuable too [35]. It is, namely, our strong conviction that the approximate methods should yield more or less uniform description of molecular properties strictly satisfying some minimal requirements. However, we shall not elaborate here the problem of designing approximate Hartree-Fock methods for molecules, because full discussion lies outside the scope of the present paper. We conclude by stating that a lot of work is necessary in this direction and by pointing out that  $\sigma^d$  and  $\chi^d$  provide simple tests *par excellence* for qualitative appraisal of atomic and molecular wavefunctions.

#### **References**

- 1. Davies DW (1967) The theory of the electric and magnetic properties of molecules. Wiley, London
- 2. Brillouin L (1934); Actual Scient. ind 71:159; Møller C, Plesset MS (1934) Phys Rev 46:618
- 3. Grimaldi F (1968) Adv. Chem Phys 14:341; Green S (1974) Adv Chem Phys 25:179
- 4. Chipman DM (1983) J. Chem Phys 78:3112; Kari RE, Csizmadia IG (1971) Theor Chim Acta 22:1
- 5. Schaeffer HF III (1972) The electronic structure of atoms and molecules. A survey of rigorous quantum mechanical results. Addison-Wesley, Reading, Mass
- 6. Jackels CF, Davidson ER (1975); J Chem Phys 63:4672
- 7. Ramsey NF (1950) Phys Rev 78:699; (1961) Am Sci 49:509
- 8. Chan SI, Dubin AS (1967) J Chem Phys 46:1745
- 9. Flygare WH Goodisman J (1968) J Chem Phys 49:3122
- 10. Deverell C (1969) Mol Pbys 17:551
- 11. Maksi6 ZB, Rupnik K (1983) Theor Chim Acta 62:397
- 12. Maksi6 ZB, Eckert-Maksi6 M, Rupnik K (1984) In: Maksi6 ZB (ed) Conceptual quantum chemistry. Models and applications. Croat chem Acta 57:1295
- 13. Maksi6 ZB, Rupnik K (1983) Z Naturforsch A 38:308
- 14. Maksi6 ZB: In: Maruani (ed) Molecules in physics, chemistry and biology. A Tribute to R. Daudel. Reidel, Dordrecht, in press
- 15. Fischer-Froese CF (1977) The Hartree-Fock method for atoms Wiley, New York
- 16. Bingham RC, Dewar MJS, Lo DH (1975) J Am Chem Soc 97:1285
- 17. Dewar MJS, Thiel W (1977) J Am Chem Soc 99:4899
- 18. Hoffmann R (1963) J Chem Phys 39:1397; (1964) 40:2480
- 19. Van der Voorn PC, Drago RS (1966); J Am Chem Soc 88:3255; White WD, Drago RS (1970) J Chem Phys 52:4717
- 20. Kovačević K, Maksić ZB (1974) J Org Chem 39:539; Maksić ZB, Rubčić A (1977) J Am Chem Soe 99:4233
- 21. Cook DB, Hollis PC, McWeeny R (1967) Mol Phys 13:553
- 22. Cook DB, McWeeny R (1968) Chem Phys Letters 1:588
- 23. Cook DB (1978) Structures and approximations for electrons in molecules. Ellis Horwood, Chichester
- 24. Clementi E, Raimondi DL (1963) J Chem Phys 38:2886
- 25. Eckert-Maksi6 M, Maksid ZB (1987) J Mol Struct Theochem 153:133 and the references cited therein
- 26. Snyder LC, Basch H (1972) Molecular wavefunctions and properties. Wiley, New York
- 27. Ray NK, Parr RG (1980) J Chem Phys 73:1334
- 28. Parr RG, Gadre SR (1980) J Chem Phys 72:3669
- 29. Burns G (1964) J Chem Phys 40:1521
- 30. Ransil BJ (1960) Rev Mod Phys 32:245
- 31. Maksi6 ZB, Bloor JE (1973) J Phys Chem 77:1520
- 32. Maksi6 ZB, Mikac N (1978) Chem Phys Letters 56:363. (1978) J Mol Struct 44:255; (1980) Mol Phys 40:455
- 33. Maksi6 ZB (in print) J Mol Struct Theochem
- 34. Siegbahn K, Nordling C, Johansson G, Hedman J, Heden PF, Hamrin K, Gelius U, Bergmark T, Werme LO, Manne R, Baer Y (1969) ESCA applied to free molecules, North-Holland, Amsterdam
- 35. Sordo JA, Pueyo L (1985) Int J Quant Chem 28:687