Acta Cryst. (1982). A38, 225–239

Pseudoatom Expansions of the First-Row Diatomic Hydride Electron Densities

BY G. S. CHANDLER AND M. A. SPACKMAN*

School of Chemistry, The University of Western Australia, Nedlands, Western Australia 6009, *Australia*

(*Received* 12 *May* 1981; *accepted* 1 *October* 1981)

Abstract

Representations of high-quality molecular electron densities are studied. An evaluation of restricted radial functions is made using a least-squares figure of merit, the molecular dipole and quadrupole moments, the electric fields at the nuclei, the electric-field gradients at the nuclei, an approximate energy and difference-density maps. For the heavy atom, a satisfactory representation has a fixed core function with a variable population and requires optimized dipolar core polarization functions, and an additional monopole term. The heavy-atom valence regions, and the H require expansions to at least the quadrupole level, with one Slater-type radial function per multipole and all exponents optimized. Additional valence radial functions and higher multipoles are required to give completely satisfactory difference-density maps but do not consistently improve the physical properties.

Introduction

It has been proposed (Stewart, 1976) that rigid pseudoatoms be used as a basis for the analysis of accurate X-ray diffraction data into static electron-density information. The static electron density is expressed as a superposition of pseudoatom densities centred on each of the nuclei of the molecule. Denoting the nuclear position vectors as $\mathbf{R}_a$ and the associated pseudoatom density as ρ_a, the total density becomes

$$\rho(\mathbf{r}) = \sum_a \rho_a(\mathbf{r} - \mathbf{R}_a) = \sum_a \rho_a(\mathbf{r}_a). \quad (1)$$

The pseudoatom densities may be expressed as finite multipole expansions about the nuclei (Dawson, 1965),

$$\rho_a(\mathbf{r}_a) = \sum_l \sum_{m=-l}^{l} \rho_{a,l}(r_a)\, P_l^m(\cos\theta_a) \begin{Bmatrix} \sin m\varphi_a \\ \cos m\varphi_a \end{Bmatrix}, \quad (2)$$

where $\rho_{a,l}(r_a)$ is a radial density function, and $P_l^m(\cos\theta_a)$ is an associated Legendre polynomial.

Kohl & Bartell (1969*a*,*b*) examined the potential of this model, for obtaining electron-density information from electron-diffraction data, by studying diatomics. Another investigation (Bentley & Stewart, 1976) on diatomic molecules made an analysis of the molecular electron-density function, concentrating on expansions up to and including quadrupole terms, $l = 2$, while restricting the $\rho_{a,l}(r_a)$ to single exponential functions. They used either standard molecular exponents (Hehre, Stewart & Pople, 1969) for the valence radial functions, or had all exponents equal to a single optimized exponent. The work concluded that single exponential radial functions reproduce qualitative features of the electron density, but do not consistently provide accurate values for physical properties. By including an extra dipole deformation term for the core, electric fields at the heavy atom were greatly improved, but other properties were left unchanged.

Previously, Stewart, Bentley & Goodman (1975) had shown that unrestricted radial functions lead to functional equations that accurately give all the molecular properties discussed by Bentley & Stewart (1976), if each pseudoatom is expanded up to the quadrupole level. Thus, a small finite multipole expansion can satisfy several static-charge properties. It is the purpose of this paper to examine the limitations of the restricted radial density functions revealed by Bentley & Stewart's (1976) study, and to show to what extent limited additions to the radial function can correct these deficiencies for the series of first-row diatomic hydrides from BH to HF.

Method

The methods used are similar to those described in a previous paper (Chandler, Spackman & Varghese, 1980). For a diatomic molecule the calculated electron density is represented as the sum of two nuclear-centred multipole expansions.

$$\rho_{\text{mol}}^c = \sum_{j=0}^{J} \rho_{a,j}(r_a)\, P_j(\cos\theta_a) + \sum_{k=0}^{K} \rho_{b,k}(r_b)\, P_k(\cos\theta_b), \quad (3)$$

* Present address: Department of Chemistry, Carnegie-Mellon University, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213, USA.

0567-7394/82/020225-15$01.00

where P_j is a Legendre polynomial and $\rho_{a,j}(r_a)$, $\rho_{b,k}(r_b)$ are trial radial density functions. If H is taken to be nucleus b, then

$$\rho_{b,k}(r_b) = \sum_{n=k}^{n_k} P_{n,k} N_{n,k} r_b^n \exp(-\zeta_n r_b). \quad (4)$$

With the heavy atom the electron density is partitioned into a spherical core and a set of valence functions. Thus the heavy-atom monopole density is represented by

$$\rho_{a,0}(r_a) = P_c \rho_c(r_a) + \sum_{n=0}^{n_0} P_{n,0} N_{n,0} r_a^n \exp(-\zeta_n r_a). \quad (5)$$

In (5) P_c is a variable core population parameter, and $\rho_c(r_a)$ is from an accurate SCF 1s atomic orbital taken from Clementi (1965). The necessity of a variable core population has been elaborated on by a number of authors (Coppens, 1977; Price, 1976; Bentley & Stewart, 1976). Higher multipoles on centre a have the same form as (4).

After substituting (5) and (4) into (3) the density $\rho^c_{\rm mol}$ is fitted to an accurate molecular electron density, $\rho_{\rm mol}$, so that the least-squares error

$$\varepsilon = \int (\rho_{\rm mol} - \rho^c_{\rm mol})^2 \, {\rm d}\tau \quad (6)$$

is minimized with respect to all populations, P_c and $P_{n,k}$ and, if required, to selected valence exponents ζ_n. The goodness of fit is measured by the least-squares figure of merit, R_w,

$$R_w = [\varepsilon / \int \rho^2_{\rm mol} \, {\rm d}\tau]^{1/2}. \quad (7)$$

Since the calculated electron density is not constrained to contain the number of electrons in the molecule, another measure of the fit is the fraction of electrons accounted for by the model, F, and given by the sum of the monopole populations divided by the number of electrons in the molecule. After the least-squares procedure, $\rho^c_{\rm mol}$ is rescaled by $1/F$. However, this rescaling is not unique and it creates difficulties which are discussed more fully later.

The choice of n in (4) and (5) for H and the notation we use for multipole functions in the following text is given in a previous paper (Chandler *et al.*, 1980). The r exponents, n, used there for the valence quadrupole, octopole and hexadecapole are also employed for the first-row atoms in this paper. With the valence monopole and dipole functions on the heavy atom, the choice of n is not so clear. A consideration of the atomic orbitals contributing to the valence density suggests that M_2 and D_2 functions would be needed. Test calculations were made on NH using a single M_0 term on H and comparing the R_w and F values obtained when single M_0, M_1, and M_2 functions were centred in turn on N. The best single monopole was M_2. Similar calculations testing single D_1 and D_2 functions were made with M_0 and D_1 terms on H and an M_2 term on N. There was little difference between the two dipole functions, and so, in anticipation of the need to employ a D_1 function for core polarization purposes, the D_2 function was chosen as the initial dipole valence function. If a further radial function is added to give a two-term expression, the power of r of the additional function is incremented by one. An earlier study on H_2 (Chandler *et al.*, 1980) found that both two- and three-term radial functions give very similar results, and so only single and two-term expansions are examined in the present paper.

Four choices for the exponents ζ_n were examined.

(1) Scheme 1. The valence exponents for the heavy atom are the sum of 2s and 2p exponents from single-ζ atomic tabulations (Clementi & Roetti, 1974). For H $\zeta = 2{\cdot}0$.

(2) Scheme 2. Exponents are twice the standard molecular value (Hehre *et al.*, 1969).

(3) Scheme 3. All valence functions on one pseudoatom share a common exponent which is optimized.

(4) Scheme 4. All valence exponents are optimized separately for each radial function.

The majority of cases studied in this paper used all multipoles up to and including quadrupole terms on both atoms. Where higher multipoles are used it is explicitly stated.

Molecular wavefunctions and physical properties

The molecules studied are the diatomic hydrides AH (A = B, C, N, O, and F) with $\rho_{\rm mol}$ determined by the near-Hartree–Fock quality wavefunctions of Cade & Huo (1967). A comparison of the physical properties calculated from $\rho^c_{\rm mol}$ and $\rho_{\rm mol}$ is made for the molecular dipole moment, μ, the quadrupole moment, Q, the electric fields at the nuclei, ε_A and $\varepsilon_{\rm H}$, and the electric-field gradient at each nucleus, q_A and q_H. For Q the origin is taken at nucleus A. The values of these properties calculated from the SCF wavefunctions are given in Table 1 (Bentley, 1974). The calculation of the properties from the pseudoatom expansion $\rho^c_{\rm mol}$ is straightforward and has been discussed by Chandler *et al.* (1980). Difference-density plots are also valuable in assessing the effectiveness of a pseudoatom expansion. The plots of $\Delta\rho_{\rm mol}$ and $\Delta\rho^c_{\rm mol}$ reported here are defined to be the difference between the molecular density and the sum of spherically averaged ground-state atoms at the same internuclear distance. This definition differs from that used by Bader, Keaveny & Cade (1967) in their discussion of the difference-density maps for the same molecules. Figs. 1 and 2 present the difference densities from $\rho_{\rm mol}$ as contour plots and as a three-dimensional (3D) graph. The 3D graphs complement the contour plots as some features are more conspicuous on the 3D graph than on the contour plot. Several trends in these maps will be important in the

comparison with $\Delta\rho^{c}_{\mathrm{mol}}$ plots. Each of the molecules shows density features typical of covalently bonded H, with an accumulation of electron density behind the A nucleus and in the bonding region surrounding the proton. The region of net accumulation of electron density around the proton contracts with increasing atomic number of the first-row atom. However, the overall peak height near the proton does not increase monotonically with increasing A atomic number but achieves its maximum at NH (Fig. 2). There is a noticeable deepening, across the period, of the hollow beyond the proton, so that the innermost negative contour increases in magnitude from −0·00195 e (a.u.)$^{-3}$ for BH [1 e (a.u.)$^{-3}$ = 6·74834 e Å^{-3}] to −0·0313 e (a.u.)$^{-3}$ for HF. Also of interest is the gradual tendency for the built-up region behind the heavy nucleus to contract and bend around towards the proton and the built-up density in the bonding region until, at HF, the two are joined. Accompanying this is a change in shape of the built-up region behind A, from a diffuse rounded area which gradually flattens and contracts until, at OH, there is a small twin peak in the plane which becomes a pair of sharp peaks either side of the F nucleus in HF. These latter features are a result of the increasing number of π-electrons as the atomic number of the heavy atom increases. Another important feature of the $\Delta\rho_{\mathrm{mol}}$ plots is the large changes in density observed near the heavy nuclei. This is in sharp contrast to the region around the H nucleus, and indicates that sharp polarization functions will be found essential in a description of the electron density on the A pseudoatom.

Table 1. *Physical properties evaluated from the Cade & Huo* (1967) *wavefunctions for the diatomic hydrides* AH

The results are taken from Bentley (1974).

Property	Molecules				
	BH	CH	NH	OH	FH
μ	0·682	0·618	0·640	0·701	0·764
Q	−2·400	−0·351	0·686	1·453	1·881
ε_A	0·006	0·009	0·010	0·009	0·008
ε_H	−0·008	−0·011	−0·019	−0·024	−0·027
q_A	−0·747	−0·943	−0·583	0·399	2·870
q_H	0·169	0·252	0·344	0·442	0·540

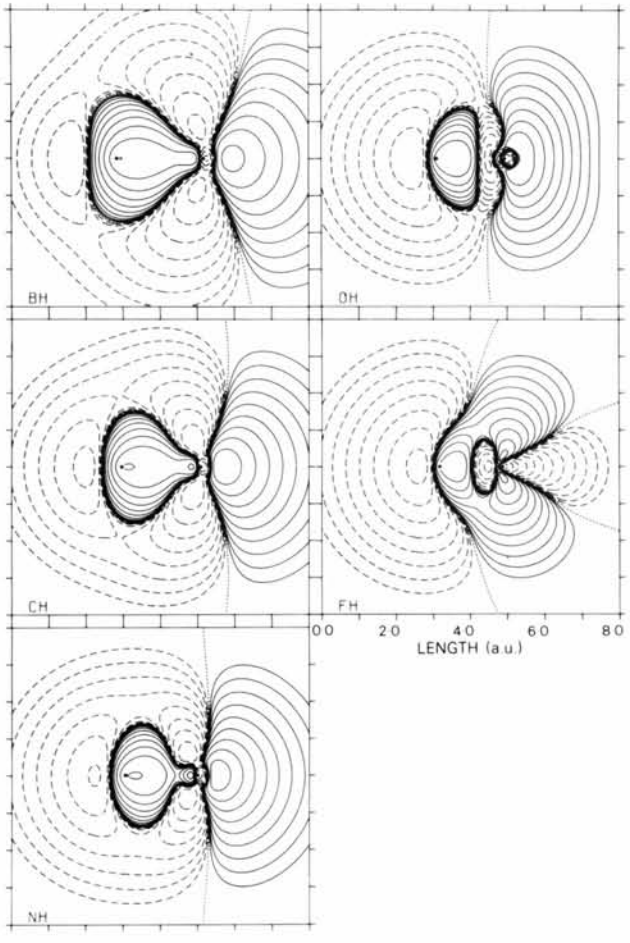

Fig. 1. Difference-density plots $\Delta\rho^{AH}_{\mathrm{mol}}$ from the wavefunctions of Cade & Huo (1967). Contour values are in e (a.u.)$^{-3}$; successive contours differ by a factor of 2; positive contours (——), negative contours (– – –), zero contours (·····). The first non-zero contours indicate an absolute value of 0·000244 e (a.u.)$^{-3}$. The heavy nucleus is indicated by a dot to the right of the centre and likewise the H to the left.

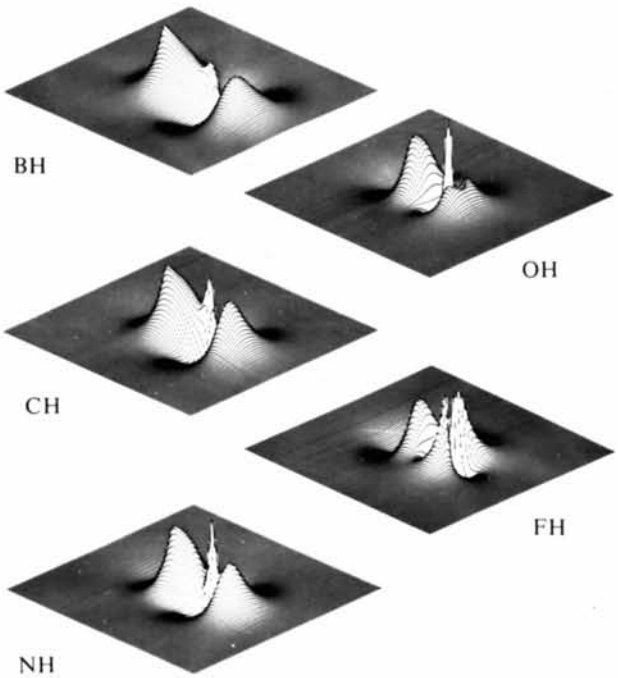

Fig. 2. Three-dimensional difference-density graphs of $\Delta\rho^{AH}_{\mathrm{mol}}$ in a plane containing both nuclei, from the wavefunctions of Cade & Huo (1967). The heavy atom is on the right in each case.

Full optimization of exponents

Although work on single-term functions in schemes 2 and 3 has already been reported (Bentley & Stewart, 1976) some aspects of the results from these schemes are commented on in the following discussion. Populations of core density functions in schemes 1 to 4 show a wide spread from 1·984 to 2·430 electrons with no obvious trends, although the populations tend to be closer to 2·0 as R_w improves. The rescaled populations of the H monopole function, supposedly a measure of the net charge transfer in the molecules, vary from −1·2 to 2·4 electrons; a result which is not chemically realistic. A two-term density function does not give any marked improvement for schemes 1 to 3 in contrast to the case of the H_2 molecule (Chandler *et al.*, 1980). When all exponents are optimized (scheme 4) there is a considerable difference between single- and two-term functions which is readily seen in the difference-density plots. Fig. 3 of the $\Delta\rho^c_{\rm mol}$ for a one-term exponential function in scheme 4 shows most of the deficiencies evident in plots from single-term functions produced with schemes 1 to 3. One conspicuous feature is the spherical excess of electron density within 1 a.u. of the heavy nucleus for all the hydrides except BH, which has a similar deficit. These features are an artifact of the rescaling procedure, and will be discussed later in the paper. Deficiencies around the heavy atom are also evident further out, where the difference density is much too contracted. However, the gross features around the proton are satisfactorily represented.

Many of the difference-density-plot discrepancies in the valence regions, especially behind the heavy nucleus, are removed on adding an extra radial function to each multipole (Fig. 4) but the regions close to the first-row atom core are still poor. This improvement with two-term functions can be attributed to the extra flexibility in the basis set so that sharp and diffuse terms combine to partially satisfy the requirement for core polarization, allowing the diffuse terms to represent the valence regions better. In this respect it is noteworthy that the set of fully optimized two-term functions are different to the others (see Tables 2 and

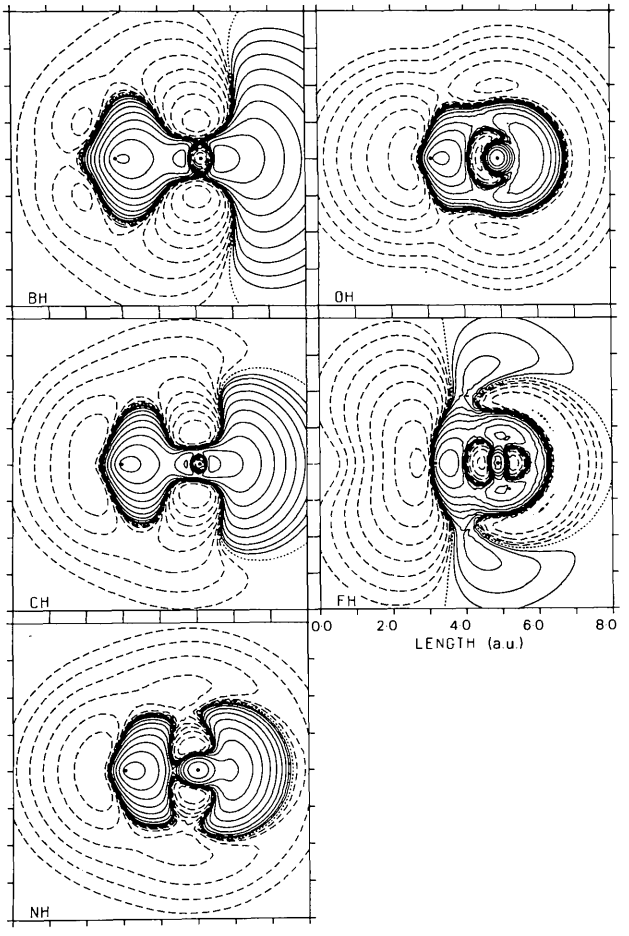

Fig. 3. Difference-density maps, $\Delta\rho^c_{\rm mol}$, for single-term functions in scheme 4. See Fig. 1 for an explanation of the contours.

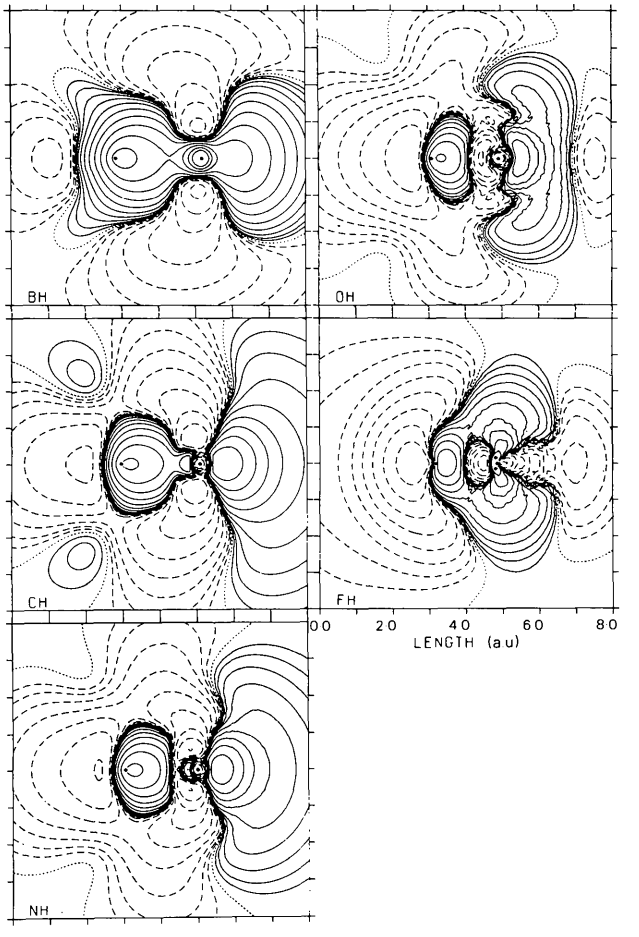

Fig. 4. Difference-density maps, $\Delta\rho^c_{\rm mol}$, for two-term functions in scheme 4. See Fig. 1 for an explanation of the contours.

3). Many of the populations for functions on H are very large, even for some dipole functions. This appears to be an indication of gross deficiencies in the basis set used on the heavy atom, which is then partially compensated for by the H-pseudoatom functions.

Tables 4 and 5 give the exponents of all valence functions under discussion. The optimized exponents in scheme 3 are identical with the analogous set reported by Bentley & Stewart (1976). From the tables, it is seen that the monopole exponents for single-term functions in scheme 4, on both A and H, are similar to those for one-term functions in scheme 3, whereas exponents of dipole and quadrupole deformation functions are appreciably different. Apparently, the optimization of exponents in scheme 3 is dominated by the monopole term on each centre, which is not surprising, in view of its large contribution compared to that of the higher order multipoles, and hence scheme 3 cannot hope to yield a suitable set of higher-order multipoles. In particular, dipole terms are usually more diffuse and quadrupole terms sharper than the corresponding monopole functions would allow. Extremely sharp quadrupole deformations are shown by OH and HF, which indicate the need for core polarization functions.

With two-term functions on the A pseudoatom in scheme 4 monopole functions generally consist of one sharp and one diffuse term, with the more diffuse term having the dominant population. Both dipole terms on A have approximately equal exponents, of the size expected for core dipole polarizations, and the quadrupole functions include sharp terms on B, N and F, but surprisingly not on O.

While two-term functions with full optimization are necessary to reproduce the valence regions of the

Table 2. *Populations for the AH series*

One-term functions in scheme 4.

		Molecule				
	Populations	BH	CH	NH	OH	HF
Heavy atom	P_c	2·003	2·054	2·099	2·128	2·104
	M_2	2·474	3·724	4·817	5·552	5·928
	D_2	−0·440	−0·249	−0·154	−0·196	−0·232
	Q_2	0·292	0·236	0·125	−0·007	−0·116
H	M_0	1·523	1·222	1·084	1·320	1·968
	D_1	0·205	0·160	0·142	0·291	0·754
	Q_2	0·066	0·051	0·034	0·058	0·339

Table 3. *Populations for the AH series*

Two-term functions in scheme 4.

		Molecule				
	Populations	BH	CH	NH	OH	HF
Heavy atom	P_c	2·199	2·047	2·061	2·058	2·099
	M_2	−0·044	3·157	3·362	4·082	4·403
	M_3	3·522	2·995	3·142	3·147	2·830
	D_2	0·066	0·077	0·074	0·070	0·062
	D_3	−0·094	−0·121	−0·126	−0·119	−0·100
	Q_2	0·435	0·074	0·001	−0·703	−0·076
	Q_3	0·001	0·129	0·090	0·569	−0·100
H	M_0	2·865	9·302	8·427	9·499	0·214
	M_1	−2·542	−10·500	−8·992	−9·786	0·454
	D_1	0·474	1·210	0·698	2·616	1·385
	D_2	−0·771	−2·029	−1·240	−3·070	−1·283
	Q_2	0·178	0·369	0·242	0·174	−0·651
	Q_3	−0·331	−0·767	−0·519	−0·458	0·630

Table 4. *Exponents of valence functions in schemes 1 to 3 for the AH series*

		Molecule				
	Exponent	BH	CH	NH	OH	HF
	ζ_A [a]	2·50	3·18	3·84	4·47	5·11
	ζ_A [b]	3·00	3·44	3·90	4·50	5·10
1-term	ζ_A	2·384	3·050	3·746	4·463	5·246
	ζ_H	2·095	2·244	2·286	2·129	1·759
2-term	ζ_A	3·463	4·127	5·041	5·756	6·197
	ζ_H	2·538	2·611	1·759	1·616	1·537

(*a*) Best-atom; (*b*) standard-molecular.

Table 5. *Optimized exponents of valence functions in scheme 4 for the AH series*

		Molecule									
		BH		CH		NH		OH		HF	
	Multipole	1-term	2-term	1-term	2-term	1-term	2-term	1-term	2-term	1-term	2-term
Heavy atom	M_2	2·435	10·387	3·073	1·872	3·747	2·352	4·474	2·921	5·272	3·900
	M_3		3·094		4·213		5·284		6·466		7·778
	D_2	2·032	6·158	2·693	7·004	3·508	8·112	3·693	9·417	3·863	11·171
	D_3		6·221		7·064		8·079		9·432		11·235
	Q_2	2·897	2·663	3·419	4·176	3·742	13·247	9·172	2·818	6·125	7·317
	Q_3		12·656		4·203		5·131		3·797		2·858
H	M_0	1·952	1·614	2·140	1·046	2·237	1·084	2·034	1·023	1·672	3·342
	M_1		1·596		1·306		1·377		1·351		3·419
	D_1	2·386	1·983	2·755	1·569	2·982	1·904	2·428	1·388	1·851	1·709
	D_2		1·969		1·687		1·840		1·728		2·109
	Q_2	2·960	2·419	3·452	2·050	4·094	2·356	3·714	2·695	2·169	0·602
	Q_3		2·399		2·073		2·205		2·270		0·746

difference density, such improvements in the density are not necessarily reflected in the physical properties derived from these expansions. Figs. 5 and 6 summarize the least-squares goodness of fit and physical-property trends for the diatomic hydrides after complete optimization of basis set exponents. These should be compared with Figs. 1 and 2 in Bentley & Stewart (1976) which show the trends for single-term functions with exponents optimized according to scheme 3. The goodness of fit for a fully optimized two-term basis set reflects the density-difference plots in that it is superior for all of the series to R_w for any of the other bases studied. This is not so for the fraction of electrons counted. The only properties for which two-term functions in scheme 4 are uniformly superior to all other bases are the electric field and the field gradient at the heavy nucleus. This is expected for ε_A, since it is the only basis containing the sharp dipole terms, which are known to be needed to reproduce ε_A (Bentley & Stewart, 1974). The results obtained for ε_H and q_H show erratic behaviour, with two-term functions giving excellent values for CH, NH and OH, but being inferior for BH and HF. In fact little is gained in optimizing all exponents when compared with scheme 3. With the quadrupole moment, a single-term basis is clearly superior. What is surprising with ε_H, q_H and Q is that best atomic exponents give excellent results, clearly superior in most cases to standard molecular exponents and also to the bases with optimized exponents. The dipole moment is poorly reproduced by all the methods used, single-term functions in scheme 3 giving about the best results except for HF.

Core deformation

The poor results from the density basis sets discussed in the preceding section are due in part to a failure to treat polarization of the core region suitably, as was recognized by Bentley & Stewart (1974). Fig. 7 displays plots of the core-deformation difference density for the first-row hydrides

$$\Delta\rho_{\text{core}} = (1\sigma_{AH})^2 - (1s_A)^2 \qquad (8)$$

where $1\sigma_{AH}$ is the 1σ orbital of Cade & Huo (1967) and $1s_A$ is the near-Hartree–Fock $1s$ atomic orbital of Clementi (1965). Deformation of the atomic $1s$ orbital is extremely localized around the heavy-atom nucleus, becomes more contracted with increasing atomic number, and is predominantly dipolar in nature. The net effect is to displace charge into the bond from behind the heavy atom and close to the nucleus. The magnitude of this polarization is exceedingly small in terms of total charge displacement, but measured in

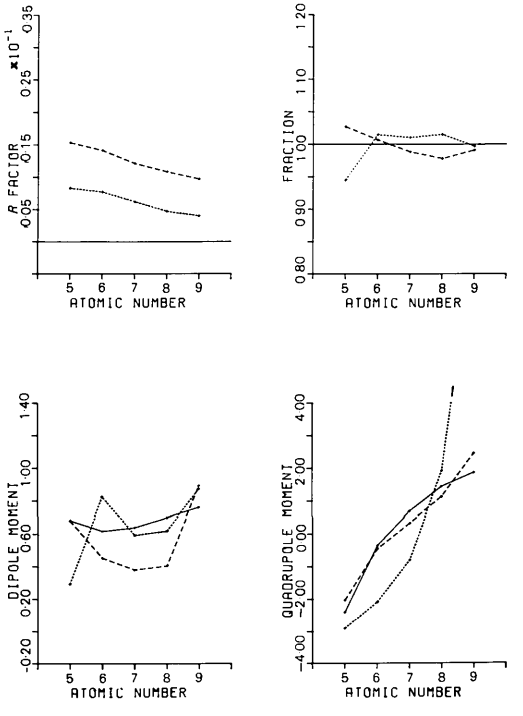

Fig. 5. Dipole moments, quadrupole moments, R_w and F for BH to HF from pseudoatom expansions with all valence exponents optimized. The SCF or ideal values are shown by a full line (——), single-term functions (– – –) and two-term functions (-----).

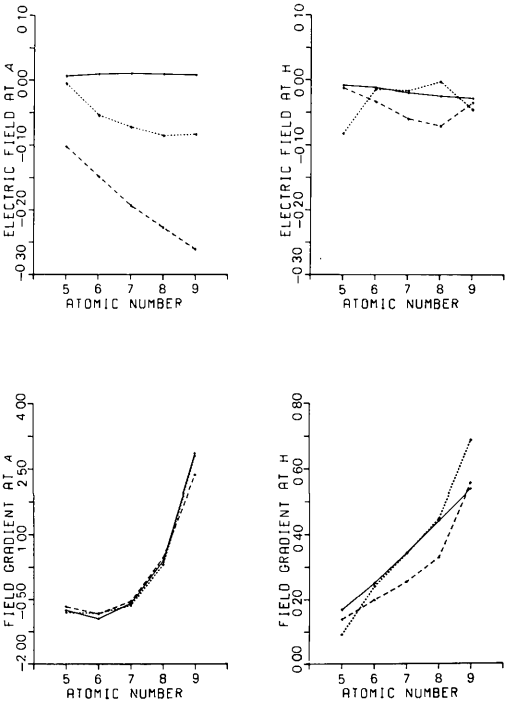

Fig. 6. Electric fields and electric-field gradients for the diatomic hydrides BH to HF, from pseudoatom expansions with all valence exponents optimized. See Fig. 5 for an explanation of the full and dashed lines.

electrostatic terms as the change in the electric field $E_{\Delta,A}$ encountered by the nucleus, above that due to the superposition of the spherical atoms, it is given by

$$E_{\Delta,A} = -\int \Delta\rho_{\text{mol}} \cos\theta r^{-2}\, \mathrm{d}^3\mathbf{r} \qquad (9)$$

(Hirshfeld & Rzotkiewicz, 1974). Because of the presence of the r^{-2} term, small asymmetries in the electron density close to the nucleus lead to large changes in the electric field. Thus, to obtain correct descriptions of the electric field at the nucleus, it is necessary to include a sharp polarization dipole and perhaps a quadrupole.

Previous studies based on the products of the lowest lying canonical 1σ orbitals (Bentley & Stewart, 1974) found quadrupolar deformation functions on all heavy atoms except F, in HF and BF. This is surprising, as other studies on the first-row hydrides indicate a steady trend of behaviour from B to F and lead to the expectation of a significant population in a sharp F quadrupole term. This term can be expected to have components arising from $1s3d$ and $2p2p$ products in higher σ molecular orbitals than the innermost 1σ. An examination of the atomic basis functions employed by Cade & Huo (1967) indicates that the most likely candidates to give a sharp Q_2 term similar to those on the other A atoms are the $(\sigma_{1s_A})(\sigma'_{3d_A})$ product, which yields Q_2 functions with exponents 5·78, 7·36, 8·57, 9·84 and 11·31 for BH to HF respectively, and the $(\sigma_{2p_A})(\sigma''_{2p_A})$ product, yielding exponents 6·53, 7·58, 8·56, 9·51 and 10·38. Of these possibilities, the $1s3d$ product is much larger than the $2p2p$ product in terms of the population of the resulting Q_2 density function. Converting the products of coefficients for the $(\sigma_{1s_A})(\sigma'_{3d_A})$ product to populations of the appropriately normalized Q_2 function results in the values in Table 6, which are presented as contributions from the individual molecular orbitals. Each molecular orbital yields quite different contributions with the 1σ orbital having by far the lowest populations and being zero in the case of HF.

Contributions from all molecular orbitals decrease with increasing atomic number, and the contributions from 2σ and 3σ molecular orbitals always have opposite signs, the 2σ contribution determining the overall sign for the resultant $1s3d$ product. This is opposite in sign to that of the 1σ contribution examined by Bentley & Stewart (1974). It is evident from Table 6, that there is likely to be a significant, sharp Q_2 term observable in a multipole analysis of FH, similar to those observed for the other AH hydrides from the 1σ molecular orbital.

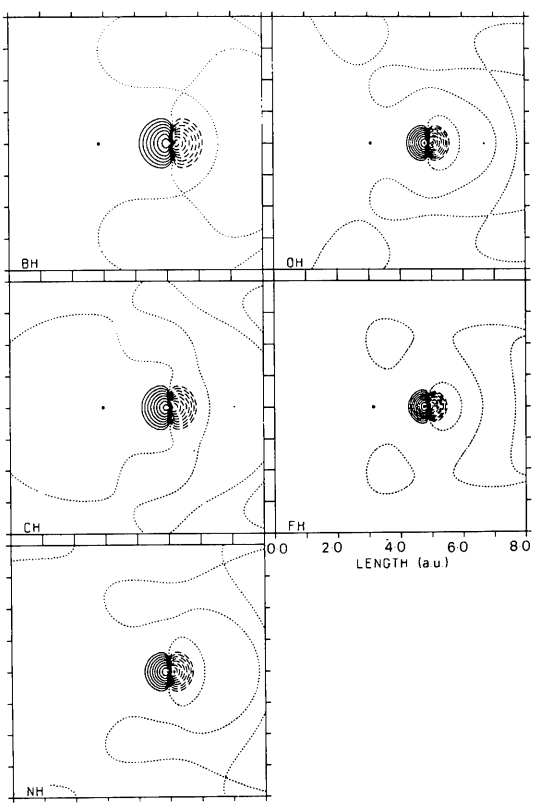

Fig. 7. Core deformation (8) in the first-row hydrides. See Fig. 1 for an explanation of the contours.

Scaling

Since our least-squares procedure is not constrained to satisfy the condition

$$\int \rho^c_{\text{mol}}\, \mathrm{d}\tau = N,$$

where N is the total number of electrons in the molecule, ρ^c_{mol} must be scaled to give the correct electron count. It has been customary in similar work to rescale ρ^c_{mol} by $1/F$, the reciprocal of the fraction of electrons counted. This method was used in the studies described earlier in this paper, but its application to both the valence and core densities produces anomalies in the region near the nucleus. A close examination of the compact spherical peaks and hollows observed in the $\Delta\rho^c_{\text{mol}}$ plots shows that they are correlated with the

Table 6. *Populations of the normalized Q_2 function resulting from the $(\sigma 1s_A)$ $(\sigma 3d'_A)$ orbital product from the wavefunctions of Cade & Huo* (1967) *for the AH series*

Populations are $\times 10^6$ electrons.

	Molecule				
	BH	CH	NH	OH	HF
Exponent	5·78	7·36	8·57	9·84	11·31
1σ contribution	119	120	72	40	0
2σ contribution	−2134	−1319	−830	−572	−398
3σ contribution	842	685	484	323	213
Total	−1173	−514	−274	−209	−185

respective F values. Large spherical peaks are shown in functions with highly scaled core populations (*i.e.* $F \ll 1$), and when F values are greater than 1·0 spherical depressions occur at the nuclei.

Evidently, the practice of rescaling the electron density by $1/F$ is responsible for these erroneous features near the nuclei. Furthermore, an examination of the relationship between the *unscaled* core populations, P_c, and the least-squares figure of merit, R_w, shown in Fig. 8, demonstrates that in general the *unscaled* core populations increase as R_w decreases and the fit to the density improves. Also it can be seen that for any R_w value the *unscaled* core population increases with atomic number. This behaviour can be anticipated, as will be discussed further in the next section, and strengthens the conclusion that any rescaling procedure should not be applied to the core populations.

Two scaling schemes which leave the core population unchanged were examined:

RS1: Only the valence monopoles are rescaled to satisfy the electron count of the molecule.

RS2: All functions except the core monopoles are rescaled.

Both scaling schemes were tested on the set of one-term functions from scheme 1. Results obtained indicated that RS2 consistently produces physical properties with a similar trend to the original scaling procedure, and the RS1 trend is usually significantly different. Plots of $\Delta\rho^c_{\rm mol}$ are improved in the vicinity of the core, both RS1 and RS2 producing almost identical plots. All further results reported in this paper are therefore based on the RS2 rescaling scheme.

Core monopole representations

A common feature of the valence density analyses developed by several authors is the assumption of a frozen core (Stewart, 1968; Stewart & Jensen, 1969; Coppens, 1971).

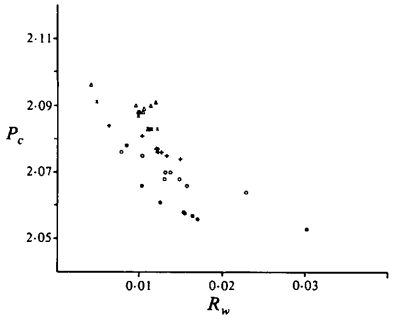

Fig. 8. Unscaled core populations as a function of the least-squares figure of merit R_w. The filled circles represent BH, open circles CH, + NH, × OH and △ HF.

Rigid adherence to a fixed core population of 2·0 in refinements has now been abandoned by many investigators in favour of a variable core population which generally fits the data better (Price, 1976; Bentley & Stewart, 1976), and the reason advanced for this is that the electron density on or near the nucleus of an isolated atom has a contribution from the $(2s)^2$ SCF orbital product as well as from the $(1s)^2$. Thus a $(1s)^2$ density populated with two electrons does not give all the charge on or near the nucleus in the isolated atom. It is useful to examine this mechanism in greater detail as it leads to an extension of the density basis set giving a much improved pseudoatom expansion for the AH molecules.

Fig. 8 indicates that the *unscaled* populations of a $(1s)^2$ core density function obtained by least-squares projection of a theoretical molecular electron density are all greater than 2·0 for B to F. The population increases as the fit to the density improves. This behaviour is most readily understood by an examination of the electron densities of the ground states of the first-row atoms where the effects of core and valence deformations do not complicate the issue.

The spherically averaged electron densities of these atoms are analysed in a manner similar to that used with molecular densities. The density basis functions employed are a $(1s)^2$ function, C, as described earlier, the normalized monopoles M_2 and M_3 and an additional M_1 function. The atomic wavefunctions are constructed from a $6s4p$ Slater basis taken from Clementi (1965).

Coefficients obtained in the analysis are rescaled according to the RS2 scheme, which in this case, with no deformation functions, is identical with RS1. Three separate basis sets with all exponents optimized were considered: (*a*) CM_2, (*b*) CM_1M_2 and (*c*) $CM_1M_2M_3$.

Tables 7 to 9 present rescaled populations, R_w and F for each of these sets. For the purposes of the present analysis the core population, P_c, can be divided into

$$P_c = 2{\cdot}0 + \delta$$

where 2·0 refers to the population required to fit the core density, $(1s)^2$, exactly and the remainder, δ, is required with the other monopole functions to fit the 'valence' density. The amount by which the core population exceeds 2·0 in Tables 7 to 9, is the factor

Table 7. *Rescaled populations and measures of fit for the basis set* (CM_2) *projected into the first-row atom densities* B *to* F (*based on RS2 rescaling*)

	Atom				
Population	B	C	N	O	H
P_c	2·064	2·073	2·081	2·088	2·094
M_2	2·936	3·927	4·919	5·912	6·906
Measures of fit					
R_w	0·0143	0·0131	0·0115	0·0107	0·0107
F	1·026	1·002	0·982	0·958	0·939

required to fit the sharp peak of the atomic 2*s* orbital at the nucleus. This peak, although small compared to that of the 1*s* orbital, is much larger than the remaining 2*s* peak in the valence region, the ratio between the two varying from 51·6 to 54·3 for B to F. This sharp peak is readily fitted by the $(1s)^2$ function, since, appropriately scaled down, it is an excellent approximation to the $(2s)^2$ function close to the nucleus. Further from the nucleus the scaled-down function overestimates the $(2s)^2$ density considerably and an M_1 function of negative sign and high exponent is required to cancel this effect. This necessity for a moderately sharp M_1 function can be simply demonstrated by considering the single-ζ wavefunction for B tabulated by Clementi & Roetti (1974). Expressing the corresponding density as a sum of monopoles gives

$$\rho_B = 2{\cdot}09M_0\,(9{\cdot}36) - 0{\cdot}18M_1\,(5{\cdot}97) + 2{\cdot}09M_2\,(2{\cdot}58) + 1{\cdot}0M_2\,(2{\cdot}42) \qquad (10)$$

where the radial exponents are in parentheses. The M_1 monopole arises naturally from the product of the χ_{1s} and χ_{2s} basis functions on squaring the 1*s* and 2*s* canonical orbitals.

From Table 7, the deficiency of the (CM_2) basis set is evident in high R_w values. The addition of an M_1 function improves the fit significantly for the lighter atoms (Table 8), but it is necessary to allow greater flexibility of the density basis set in the outer regions of the heavier atoms, and hence addition of an M_3 function is essential (Table 9), and an excellent fit is then obtained for all atoms. Fig. 9 illustrates the resulting valence radial densities for the B atom, rescaled as in the RS2 scheme. The (CM_2) set is quite inadequate, but both the (CM_1M_2) and $(CM_1M_2M_3)$ sets fit the electron density well; the former is rescaled more, causing it to appear a poor fit in the figure. The other first-row atoms, C to F, show similar behaviour.

In the first-row atoms only the 1*s* and 2*s* orbitals contribute to the electron density at the nucleus. If the least-squares fit to this peak at the nucleus is exact, and is made up only from the $(1s)^2$ core function, then the excess of P_c above 2·0 would be given by

$$2\rho_{2s}^{SCF}(0)/\rho_{1s}^{SCF}(0),$$

where $\rho_{1s}^{SCF}(0)$ and $\rho_{2s}^{SCF}(0)$ are the densities at the nucleus of the 1*s* and 2*s* orbitals (normalized to 1·0 electron). From the atoms B to F, this ratio from the Clementi (1965) wavefunctions is 0·082, 0·091, 0·097, 0·103 and 0·107. The values of P_c for the best atomic density fits, $(CM_1M_2M_3)$ in Table 9, all exceed 2·0 by more than the ratios given above, indicating that the electron density at the nucleus is overestimated in all cases, probably due to remaining small deficiencies in the density basis set.

Atomic studies, then, indicate that addition of an M_1 monopole to the heavy-atom density basis set will improve the valence density model for the diatomic hydrides. Inclusion of this term into the least-squares analysis should allow the other functions in the heavy-atom basis greater flexibility to describe the outer valence regions of the pseudoatom density.

Table 8. *Rescaled populations and measures of fit for the basis set* (CM_1M_2) *projected into the first-row atom densities* B *to* F (*based on RS2 rescaling*)

	Atom				
Population	B	C	N	O	F
P_c	2·087	2·098	2·105	2·112	2·118
M_1	−0·204	−0·152	−0·108	−0·074	−0·051
M_2	3·117	4·054	5·003	5·962	6·933
Measures of fit					
R_w	0·0022	0·0032	0·0045	0·0066	0·0087
F	0·975	0·967	0·959	0·944	0·932

Table 9. *Rescaled populations and measures of fit for the basis set* $(CM_1M_2M_3)$ *projected into the first-row atom densities* B *to* F (*based on RS2 rescaling*)

	Atom				
Population	B	C	N	O	F
P_c	2·086	2·095	2·102	2·108	2·112
M_1	−0·830	−0·836	−0·776	−0·788	−0·631
M_2	2·648	3·202	3·720	4·276	4·693
M_3	1·097	1·538	1·954	2·405	2·826
Measures of fit					
R_w	0·0007	0·0007	0·0007	0·0009	0·0008
F	1·001	1·000	0·999	0·997	0·995

Density fits with core polarization and M_1 terms included

The effect of two core polarization terms, a dipole, and a quadrupole have been examined. Addition of these functions, abbreviated as D_1^* and Q_2^*, to the set of

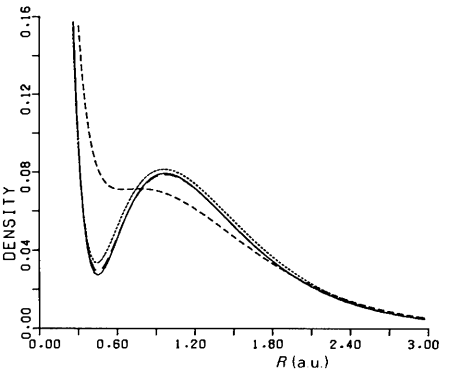

Fig. 9. Monopole expansions of the radial density [in e (a.u.)$^{-3}$] for the B atom. The SCF density is given by the unbroken line (——), CM_2 expansion (– – –), CM_1M_2 expansion (- - - - -) and the $CM_1M_2M_3$ expansion (— — — —) which is almost entirely coincident with the SCF curve.

one-term functions gives a basis consisting of a core function, C, one monopole, M_2, two dipole, D_1 and D_2, and two quadrupole functions, Q_2^* and Q_2, on the heavy atom, and a monopole, M_0, a dipole, D_1, and a quadrupole function on H; a basis which will be abbreviated as $(CM_2D_1^*D_2Q_2^*Q_2|M_0D_1Q_2)$. This was assessed using scheme 4, and the RS2 scaling procedure. There is no marked improvement in R_w on inclusion of the polarization terms and most of the physical properties do not change greatly, but ε_A and q_A values improve markedly. Difference-density plots $\Delta\rho^c_{\text{mol}}$ are given in Fig. 10. Comparisons with the analogous plots in Fig. 3 from a $(CM_2D_2Q_2|M_0D_1Q_2)$ basis with the earlier overall rescaling shows that the effect of adding core polarization functions is not very noticeable outside the core region. There is a reorganization of charge, most noticeable in the heavier elements of the series, leading to a charge concentration along the bond direction. Where this was already present it is further developed, and in the other cases it is a new feature. However, in HF there is a new feature in the appearance of the distorted π-like density just above and below the internuclear axis. In most cases though there is still an excess of charge, compared to the reference plots, Fig. 1, surrounding the heavy atom at moderate distances from the nucleus. This can be attributed to the absence of an M_1 function.

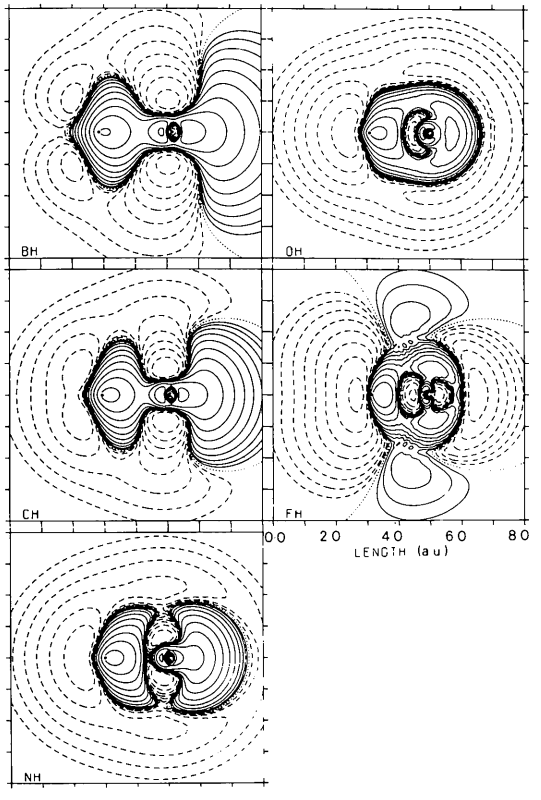

Fig. 10. Difference-density plots for a $(CM_2D_1^*D_2Q_2^*Q_2|M_0D_1Q_2)$ basis. For an explanation of the contours see Fig. 1.

When an M_1 term is added to give a $(CM_1M_2D_1^*D_2Q_2^*Q_2|M_0D_1Q_2)$ basis and all exponents are reoptimized the physical properties are greatly improved (Table 10). Table 11 contains the corresponding optimized exponents. It shows the comparatively high exponents for the M_1 functions, lying between those of the core and valence functions. With their small negative populations, these functions remove the excess electron density surrounding the heavy nuclei. A large improvement in R_w is obtained so that all values are below 0·0045. Coupled with this is an increase in the core populations of approximately 0·02 electrons. Physical properties are now more consistently close to the SCF values. Dipole moments for CH, NH, OH and HF are very close to the SCF values; Q and ε_A are well reproduced for BH and CH, but the agreement worsens as the atomic number increases, and ε_H, q_A and q_H are well reproduced in all cases.

Table 10. *Physical properties, populations and measures of fit (based on RS2 rescaling) of the AH hydrides for the basis set* $(CM_1M_2D_1^*D_2Q_2^*Q_2|M_0D_1Q_2)$ *with all exponents optimized*

		Molecule				
	Population	BH	CH	NH	OH	HF
Heavy atom	P_c	2·078	2·089	2·096	2·103	2·108
	M_1	−0·477	−0·327	−0·222	−0·147	−0·087
	M_2	2·184	3·032	3·903	4·628	5·432
	D_1^*	0·007	0·006	0·005	0·004	0·004
	D_2	−0·734	−0·618	−0·512	−0·477	−0·408
	Q_2^*	0·050	0·060	−0·044	2·150	−0·039
	Q_2	0·166	0·083	0·091	−2·261	−0·190
H	M_0	2·214	2·206	2·222	2·417	2·547
	D_1	0·405	0·526	0·654	0·813	0·950
	Q_2	0·101	0·156	0·239	0·322	0·437
	Property					
	R_w	0·0041	0·0040	0·0040	0·0041	0·0045
	F	0·996	0·996	0·997	1·007	1·014
	μ	0·765	0·610	0·640	0·701	0·766
	Q	−2·555	−0·344	1·080	2·201	2·384
	ε_A	0·009	0·026	0·043	0·058	0·071
	ε_H	−0·027	−0·033	−0·033	−0·021	−0·008
	q_A	−0·752	−0·963	−0·463	0·114	2·934
	q_H	0·132	0·212	0·320	0·443	0·585

Table 11. *Optimized exponents for the basis set* $(CM_1M_2D_1^*D_2Q_2^*Q_2|M_0D_1Q_2)$

		Molecule				
	Exponent	BH	CH	NH	OH	HF
Heavy atom	M_1	4·053	5·576	7·374	9·622	12·992
	M_2	3·078	3·612	4·201	4·841	5·476
	D_1^*	8·659	11·591	14·414	17·475	20·220
	D_2	1·956	2·370	2·772	3·051	3·367
	Q_2^*	3·214	4·062	7·226	3·209	8·670
	Q_2	3·291	4·162	6·162	3·184	3·542
H	M_0	1·714	1·726	1·712	1·653	1·599
	D_1	1·957	1·956	1·927	1·920	1·920
	Q_2	2·626	2·578	2·425	2·385	2·287

From the $\Delta\rho^c_{mol}$ plots in Fig. 11 it can be seen that the core region has been improved, the valence regions are also better, and overall the comparison with the accurate difference densities in Fig. 1 is now favourable although the agreement worsens as the atomic number increases.

It was noted earlier (Fig. 4) that the heavier atoms demand the extra flexibility afforded by two-term functions before their difference-density plots resemble reference examples (Fig. 1). Accordingly, the basis set was further expanded to contain two-term valence functions. Least-squares correlation difficulties were encountered with the addition of a Q_3 term on the first-row atom and the valence Q_2 term was dropped leaving the basis $(CM_1M_2M_3D_1^*D_2D_3Q_2Q_3|M_0M_1D_1D_2Q_2Q_3)$. The asterisk is dropped from the Q_2 term as it is no longer strictly a core polarization function. Again all exponents were reoptimized. The resulting least-squares parameters, populations and physical properties are given in Table 12 and the corresponding exponents in Table 13. From Table 12 it is seen that the R_w values have again improved, so that they are now in the range obtained by Bentley & Stewart (1975) from generalized scattering factors. Physical properties are again consistently reproduced giving satisfactory agreement with the SCF values. Nevertheless, they are often in worse agreement than those from the previous single-term basis and they are often sufficiently far from the respective SCF values to make it doubtful whether it is possible reliably to reproduce physical properties from restricted radial functions. A plot of $\Delta\rho^c_{mol}$ Fig. 12 now agrees with Fig. 1 in all the major features. Remaining discrepancies, however, indicate the need for an octopole function on the H especially in

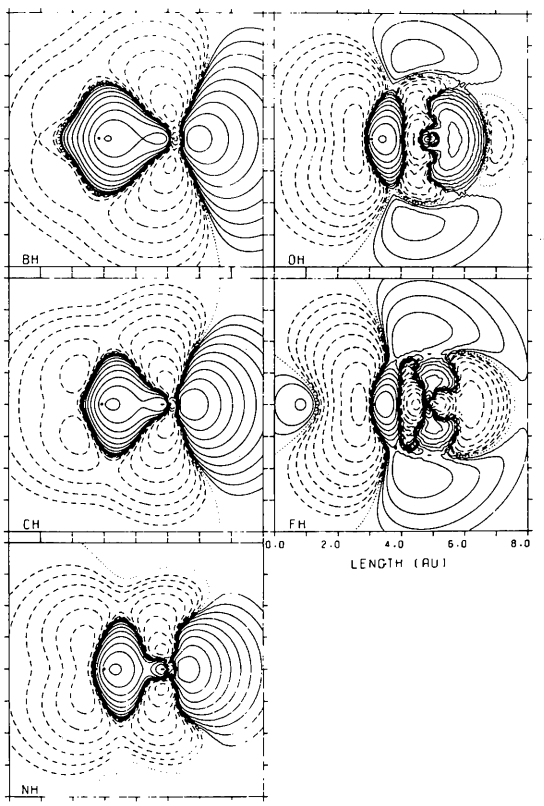

Fig. 11. Difference-density plots for a $(CM_1M_2D_1^*D_2Q_2^*Q_2|M_0D_1Q_2)$ basis. For an explanation of the contours see Fig. 1.

Table 12. *Physical properties, populations and measures of fit, based on RS2 rescaling, of the AH hydrides for the basis set* $(CM_1M_2M_3D_1^*D_2D_3Q_2Q_3|M_0M_1D_1D_2Q_2Q_3)$ *with all exponents optimized*

		Molecule				
	Population	BH	CH	NH	OH	HF
Heavy atom	P_c	2·078	2·089	2·096	2·102	2·106
	M_1	−0·626	−0·498	−0·403	−0·532	−0·443
	M_2	2·481	5·009	7·092	12·637	15·656
	M_3	−0·426	−1·881	−2·773	−6·853	−8·804
	D_1^*	0·006	0·018	0·017	0·013	0·012
	D_2	−0·605	−0·716	−0·555	−0·386	−0·306
	D_3	−0·259	−0·024	−0·039	−0·051	−0·054
	Q_2	0·371	0·257	0·190	−0·006	−0·073
	Q_3	−0·266	−0·230	−0·260	−0·161	−0·191
H	M_0	2·346	1·547	1·222	0·861	0·767
	M_1	0·148	0·733	0·766	0·785	0·718
	D_1	0·225	0·775	0·756	0·737	1·088
	D_2	0·207	−0·331	−0·341	−0·389	−0·738
	Q_2	−0·066	0·450	0·471	0·629	0·433
	Q_3	0·142	−0·338	−0·357	−0·526	−0·340
	Property					
	R_w	0·0032	0·0031	0·0026	0·0024	0·0018
	F	1·005	1·008	1·007	1·004	1·003
	μ	0·676	0·530	0·547	0·573	0·743
	Q	−2·505	−1·115	0·049	1·024	1·571
	ε_A	0·010	0·004	0·001	0·001	−0·002
	ε_H	0·003	−0·016	−0·020	−0·026	−0·029
	q_A	−0·706	−0·862	−0·536	0·501	2·798
	q_H	0·200	0·219	0·306	0·397	0·500

Table 13. *Optimized exponents for the basis set* $(CM_1M_2M_3D_1^*D_2D_3Q_2Q_3|M_0M_1D_1D_2Q_2Q_3)$ *for the AH series*

		Molecule				
	Population	BH	CH	NH	OH	HF
Heavy atom	M_1	3·736	4·908	6·165	6·592	7·794
	M_2	3·065	3·386	3·936	4·345	4·880
	M_3	2·481	3·557	4·389	4·987	5·661
	D_1^*	9·012	7·828	9·061	11·114	12·473
	D_2	1·996	2·170	2·374	2·492	2·483
	D_3	2·108	8·160	8·465	8·622	9·489
	Q_2	2·829	3·434	3·459	9·788	7·398
	Q_3	2·484	2·645	2·944	2·507	3·236
H	M_0	1·677	1·932	2·079	2·286	2·332
	M_1	1·675	1·935	2·147	2·300	2·351
	D_1	1·763	1·762	1·869	1·948	1·799
	D_2	2·845	1·686	1·806	1·944	1·966
	Q_2	3·001	2·097	2·223	2·208	2·431
	Q_3	3·672	2·343	2·507	2·530	2·744

BH, CH and NH. When an octopole function is added to the single-term basis for each pseudoatom giving $(CM_1M_2D_1^*D_2Q_2Q_3O_3|M_0D_1Q_2O_3)$, and all exponents are reoptimized, there is an improvement in difference density plots, Fig. 13. It is most evident around the H atom in the lower hydrides where the bulge in the electron-density contours, along the direction of the bond axis, and beyond the proton, is removed and the depression, also beyond the proton, conforms more closely to the reference plots, Fig. 1. The changes for OH and HF are not so marked and the plots still have deficiencies around the heavy atom which, as has already been pointed out, seem to require at least two-term functions to remedy them.

Least-squares populations, parameters and physical properties obtained with this last basis set are given in Table 14 and the corresponding optimized exponents are in Table 15. Although the R_w factors are all lower, with the greatest lowering occurring in BH, and the improvement decreasing with increasing atomic number, there is no consistent improvement in the values of the physical properties when compared to the results from the analogous basis without the octopoles (Table 10). Dipole and quadrupole moments are in slightly worse agreement with the SCF values, but the electric fields and field gradients at the nuclei are generally better.

Table 14. *Physical properties, populations and measures of fit (based on RS2 rescaling) for the basis set* $(CM_1M_2D_1^*D_2Q_2Q_3O_3|M_0D_1Q_2O_3)$ *with all exponents optimized for the AH series*

		Molecule				
	Population	BH	CH	NH	OH	HF
Heavy atom	P_c	2·078	2·089	2·096	2·102	2·107
	M_1	−0·347	−0·382	−0·299	−0·181	−0·106
	M_2	2·314	2·868	3·593	4·358	5·174
	D_1^*	0·007	0·007	0·006	0·005	0·004
	D_2	−0·605	−0·627	−0·527	−0·405	−0·305
	Q_2	0·266	0·038	−0·085	−0·279	−0·051
	Q_3	†	0·097	0·092	0·199	−0·145
	O_3	0·061	0·036	0·025	0·021	0·014
H	M_0	1·954	2·425	2·609	2·721	2·825
	D_1	0·332	0·685	0·905	1·083	1·219
	Q_2	0·087	0·273	0·446	0·593	0·719
	O_3	0·015	0·063	0·146	0·226	0·310
	Property					
	R_w	0·0017	0·0019	0·0021	0·0025	0·0032
	F	0·990	0·993	1·001	1·006	1·016
	μ	0·674	0·709	0·775	0·913	1·044
	Q	−2·386	0·177	1·427	2·177	2·442
	ε_A	0·016	0·024	0·032	0·038	0·039
	ε_H	−0·025	−0·032	−0·031	−0·025	−0·004
	q_A	−0·722	−0·868	−0·435	0·472	2·862
	q_H	0·146	0·227	0·329	0·440	0·568

† Q_3 on B refined to a very sharp function with very small population; it was therefore dropped from the basis set.

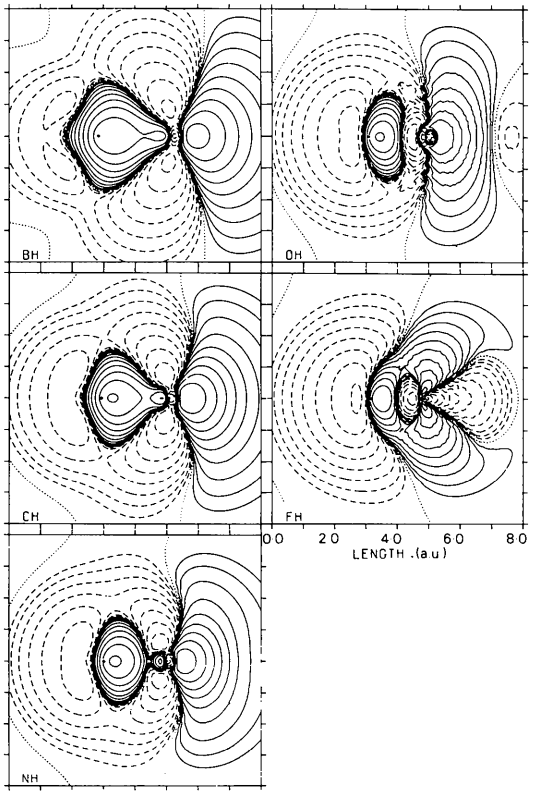

Fig. 12. Difference-density plots for a $(CM_1M_2M_3D_1^*D_2D_3Q_2Q_3|M_0M_1D_1D_2Q_2Q_3)$ basis. For an explanation of the contours see Fig. 1.

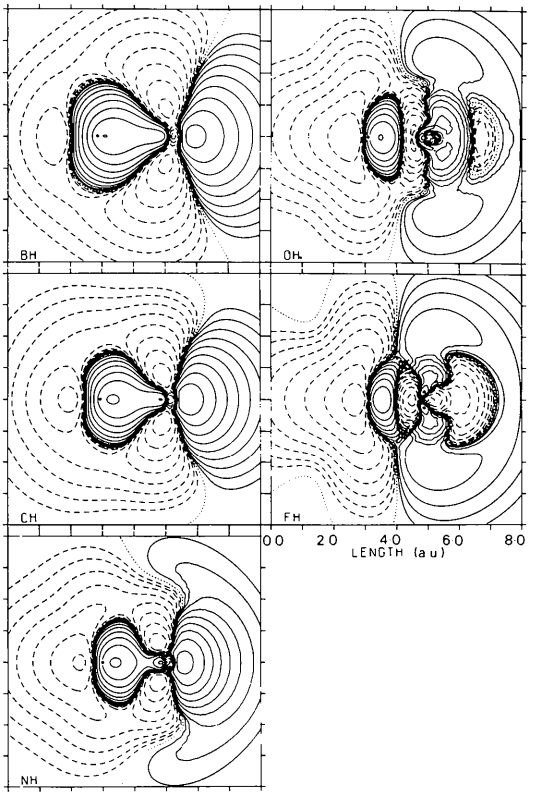

Fig. 13. Difference-density plots for a $(CM_1M_2D_1^*D_2Q_2Q_3O_3|M_0D_1Q_2O_3)$ basis. For an explanation of the contours see Fig. 1.

The electronic energy

Recently, approximate energy relationships for molecules which are suited to applications with a pseudoatom expansion have been published (Politzer, 1976, 1979). Politzer (1976) investigated the expression

$$E_{\text{mol}} = \tfrac{3}{7} \sum_{\text{atoms, } A} Z_A V_A^{\text{mol}}, \qquad (11)$$

which relates the total energy of the molecule to the sum over the constituent nuclei of the nuclear charge multiplied by the total electrostatic potential at each nucleus (V_A^{mol}). With V_A^{mol} values from SCF wavefunctions this relationship reproduces the SCF energies to within 2%. Bentley (1979) has used (11) to calculate electronic energies from experimentally determined electron densities. Politzer (1979) has subsequently improved this relationship by allowing the constant $\frac{3}{7}$ to become a function of atomic number,

$$E_{\text{mol}} = \sum_A k_A Z_A V_A^{\text{mol}}. \qquad (12)$$

Values of k_A are determined from application of (12) to numerical Hartree–Fock calculations and are tabulated by Politzer (1979) for some elements up to Bi. Using the exact E_{atom} and $V_{\text{H}}^{\text{atom}}$ for H gives $k_{\text{H}} = 0{\cdot}5$. Politzer questions the use of a value obtained for a one-electron system in treating molecules containing many electrons. Therefore, he has calculated k_{H} for the first-row diatomic hydrides from the wavefunctions employed in the present paper (Cade & Huo, 1967), and atomic k_A values. The resulting k_{H} values are quite different from 0·5 except for OH. This difference has very little effect on the calculated energy however, as $V_{\text{H}}^{\text{mol}}$ is invariably close to $-1{\cdot}0$ and with $Z_{\text{H}} = 1{\cdot}0$, the contribution from the H nucleus to E_{mol} in (12) is understandably small. Calculating E_{mol} *via* (12) for a variety of wavefunctions, including some involving configuration interaction, reproduces the SCF energy to within an average deviation of 0·2%, regardless of whether k_{H} is 0·5 or determined from the hydrides (Politzer, 1979).

The expression (12) has been applied to all of the density function expansions discussed in this paper after rescaling with the RS2 scheme. The k_{H} appropriate to each AH molecule was employed. These approximate energies are displayed in Table 16 as values of the percentage deviation (%D) from the SCF energies.

Values of %D as a function of F, the fraction of electrons accounted for by the model are displayed in Fig. 14 and as a function of R_w in Fig. 15. The calculated energy is not a function of R_w, but, with few

Table 15. *Optimized exponents for the basis set* $(CM_1M_2D_1^*D_2Q_2Q_3O_3 \mid M_0D_1Q_2O_3)$, *for the* AH *series*

		Molecule				
	Population	BH	CH	NH	OH	HF
Heavy atom	M_1	4·369	5·354	6·796	9·110	12·258
	M_2	2·912	3·692	4·342	4·936	5·552
	D_1^*	9·098	11·264	13·770	16·219	18·546
	D_2	1·997	2·456	2·899	3·410	3·938
	Q_2	3·037	4·830	3·204	4·157	8·166
	Q_3	†	4·852	5·659	5·594	4·519
	O_3	3·243	4·223	5·153	5·842	6·782
H	M_0	1·786	1·654	1·606	1·549	1·493
	D_1	1·991	1·793	1·774	1·762	1·763
	Q_2	2·499	2·166	2·088	2·076	2·077
	O_3	3·427	2·845	2·428	2·324	2·246

† Q_3 on B refined to a very sharp function with very small population; it was therefore dropped from the basis set.

Table 16. *The percentage deviation between the SCF energies and energies obtained from pseudoatom expansions of the density via equation* (12)

	%D				
Expansion†	BH	CH	NH	OH	HF
(*a*)	−1·05	−1·03	−0·90	−0·73	−0·76
(*b*)	0·17	0·10	−0·37	−0·63	−0·68
(*c*)	−4·24	−3·11	−1·96	−1·69	−1·51
(*d*)	−0·45	−0·25	−0·36	−0·73	−1·18
(*e*)	−0·50	−0·61	−0·82	−0·95	−0·74
(*f*)	−1·14	−1·20	−1·27	−0·64	0·01
(*g*)	−0·05	−0·24	−0·50	−0·53	−0·13
(*h*)	−1·37	0·27	0·32	0·76	0·09
(*i*)	−0·14	−0·34	−0·59	−0·73	0·11
(*j*)	−0·22	−0·11	−0·03	0·43	0·68
(*k*)	0·00	0·23	0·23	0·35	0·29
(*l*)	−0·39	−0·20	0·06	–	–

† Key to the expansions, the number following the basis indicates the scheme used for exponent determination

(*a*) 1-term, 1; (*b*) 2-term, 1; (*c*) 1-term, 2;
(*d*) 2-term, 2; (*e*) 1-term, 3; (*f*) 2-term, 3;
(*g*) 1-term, 4; (*h*) 2-term, 4;
(*i*) $(CM_2D_1^*D_2Q_2^*Q_2 \mid M_0D_1Q_2)$, 4;
(*j*) $(CM_1M_2D_1^*D_2Q_2^*Q_2 \mid M_0D_1D_2Q_2Q_3)$, 4;
(*k*) $(CM_1M_2M_3D_1^*D_2D_3Q_2Q_3 \mid M_0M_1D_1D_2Q_2Q_3)$, 4;
(*l*) $(CM_1M_2D_1^*D_2Q_2Q_3O_3 \mid M_0D_1Q_2Q_3)$, 4.

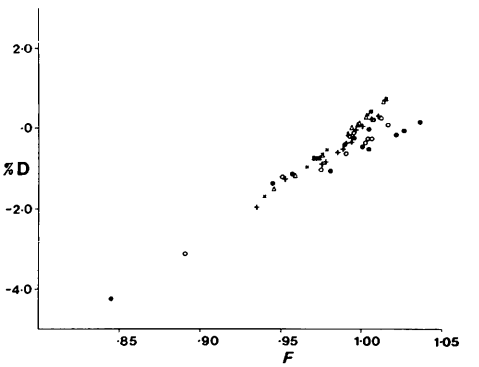

Fig. 14. The percentage deviation of the approximate energy (12) from the SCF energy, as a function of the fraction of electrons counted. See Fig. 8 for code.

exceptions, the points in Fig. 15 are grouped in two areas. One, containing the majority of the points, shows energies more negative than E_{mol}^{SCF}, with moderately large R_w values. The other group contains all the expansions with an M_1 term in the basis set, plus a few others, clustered in a region largely of positive %D, and very low R_w. There is also a trend observable in this cluster; BH and CH have %D values close to zero or small negative, while OH and FH yield %D values greater than zero. This behaviour is seen in Fig. 14 to be due to the large F values for these expansions. The deviation of the calculated energy about E_{mol}^{SCF} is an almost linear function of F, with the various hydride points spread about a line passing through $F = 1{\cdot}0$, $\%D = 0{\cdot}0$ so that for F close to 1·0, $E_{mol}^{c} \approx E_{mol}^{SCF}$.

Table 16 contains the %D values obtained for the hydrides. The average %D for each molecule is 0·81, 0·64, 0·62, 0·72 and 0·58 for BH to HF respectively. Overall this represents an average %D of 0·67% which is comparable with the value of 0·2% obtained by Politzer (1979), and this figure includes particularly bad expansions such as those for one-term functions in scheme 2, (*c*), in Table 16. Scaling plays an important role in determining the accuracy of reproduction of E_{mol}^{SCF}. For example, scaling ρ_{mol}^{c} by $1/F$ yields an average %D of 2·19% and surprisingly the unscaled ρ_{mol}^{c} yields an average of 0·43% for %D. The large average percent deviation observed for densities rescaled by $1/F$ is a further indication of the erroneous nature of this procedure. It is so evident here because of the dominance of the core electrons in determining the electrostatic potential at the nucleus. The contribution of a $(1s)^2$ core function populated by 2·0 electrons, to the atomic energy, calculated *via* (12) is between 65% and 82% of E_{atom} for the atoms B to F.

Conclusions

Simple extensions improve the accuracy of pseudoatom expansions of electron densities from accurate SCF wavefunctions for the first-row diatomic hydrides over those obtained by Bentley & Stewart (1976). These extensions have important implications for electron-density studies of experimental data. Some indefiniteness remains, however, in incorporating the conclusions into the treatment of experimental data, because no reciprocal-space analysis has been performed to gauge the effect of dealing with a finite Ewald sphere of data. Thus, the inclusion of sharp core polarization functions into the density basis was found to be necessary to improve the electric field and electric-field gradients at the heavy nuclei, but, because of their contracted nature, it is unlikely to be important in the experimental analysis. However, it is clear from our study that core populations need to be allowed to vary and not be held fixed at 2·0. Also the core population should not be included in rescaling designed to give the correct total electron count. This correction should only be applied to the valence populations. When the core population is rescaled spuriously large peaks and hollows appear around the heavy nucleus in the difference densities and, as a consequence, energy values calculated using Politzer's (1979) approximate relation (12) are poor. Another factor of major importance is the inclusion of an M_1 function with an exponent midway between a core and valence function in the pseudoatom basis. This was needed in order to achieve consistent agreement of physical properties with the corresponding SCF values, and would play a similar role in experimental data analysis.

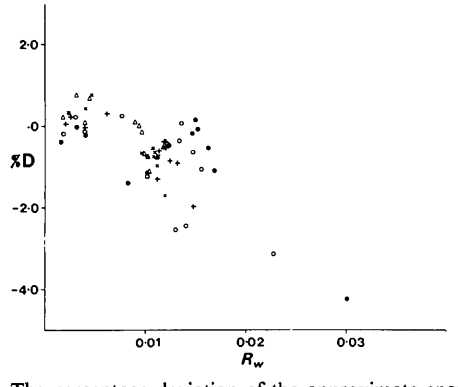

Fig. 15. The percentage deviation of the approximate energy (12) from the SCF energy, as a function of R_w. See Fig. 8 for code.

Further additions to the basis set such as two-term valence multipoles or the inclusion of an octopole function do not bring about any consistent or sizeable improvements in the physical properties. However, these further refinements do make favourable changes to difference-density plots. Two-term valence functions are needed to obtain the correct difference-density features around F and O in OH and HF, and octopole functions are needed to refine the features around the H in BH, CH and NH.

A disturbing feature of this study is the extreme variability to basis-set change shown by total electron populations on each atom (Tables 2, 3, 10, 12, 14) as measured by the sum of the monopole populations on each centre. It indicates that the present procedures are a shaky foundation on which to base a population analysis. Some comment ought also to be made about the stability of physical properties to density basis-set changes. The best agreement obtained here was for a single-ζ basis set with core polarization functions, and an added M_1 function $(CM_1M_2D_1^*D_2Q_2^*Q_2|M_0D_1Q_2)$. When either extra valence basis functions or a higher multipole in the form of an octopole are added there is an overall deterioration in the property values, even though R_w values and the difference-density plots improve considerably. This erratic behaviour indicates that caution should be exercised when using a pseudoatom expansion to estimate physical properties.

MAS gratefully acknowledges the support of a Commonwealth Postgraduate Research Award for the duration of this work.

We wish to thank the Australian Research Grants Commission for support.

References

BADER, R. F. W., KEAVENY, I. & CADE, P. E. (1967). *J. Chem. Phys.* **47**, 3381–3402.
BENTLEY, J. J. (1974). *Charge Density Analysis of Coherent X-ray Scattering by Diatomic Molecules.* Thesis, Carnegie-Mellon Univ.
BENTLEY, J. J. (1979). *J. Chem. Phys.* **70**, 159–164.
BENTLEY, J. J. & STEWART, R. F. (1974). *Acta Cryst.* A**30**, 60–67.
BENTLEY, J. J. & STEWART, R. F. (1975). *J. Chem. Phys.* **63**, 3794–3803.
BENTLEY, J. J. & STEWART, R. F. (1976). *Acta Cryst.* A**32**, 910–914.
CADE, P. E. & HUO, W. M. (1967). *J. Chem. Phys.* **47**, 614–648.
CHANDLER, G. S., SPACKMAN, M. A. & VARGHESE, J. N. (1980). *Acta Cryst.* A**36**, 657–669.
CLEMENTI, E. (1965). *Tables of Atomic Functions.* Suppl. to *IBM J. Res. Dev.* **9**, 2.
CLEMENTI, E. & ROETTI, C. (1974). *At. Data Nucl. Data Tables*, **14**, 177–478.
COPPENS, P. (1971). *Acta Cryst.* B**27**, 1931–1938.
COPPENS, P. (1977). *Isr. J. Chem.* **16**, 159–162.
DAWSON, B. (1965). *Aust. J. Chem.* **18**, 595–603.
HEHRE, W. J., STEWART, R. F. & POPLE, J. A. (1969). *J. Chem. Phys.* **51**, 2657–2664.
HIRSHFELD, F. L. & RZOTKIEWICZ, S. (1974). *Mol. Phys.* **27**, 1319–1343.
KOHL, D. A. & BARTELL, L. S. (1969*a*). *J. Chem. Phys.* **51**, 2891–2895.
KOHL, D. A. & BARTELL, L. S. (1969*b*). *J. Chem. Phys.* **51**, 2896–2904.
POLITZER, P. (1976). *J. Chem. Phys.* **64**, 4239–4240.
POLITZER, P. (1979). *J. Chem. Phys.* **70**, 1067–1069.
PRICE, P. F. (1976). *The Electron Density in Molecular Crystals.* Thesis, Univ. of Western Australia.
STEWART, R. F. (1968). *J. Chem. Phys.* **48**, 4882–4889.
STEWART, R. F. (1976). *Acta Cryst.* A**32**, 565–574.
STEWART, R. F., BENTLEY, J. & GOODMAN, B. (1975). *J. Chem. Phys.* **63**, 3786–3793.
STEWART, R. F. & JENSEN, L. H. (1969). *Z. Kristallogr.* **128**, 133–147.

Acta Cryst. (1982). A**38**, 239–247

A Procedure for Joint Refinement of Macromolecular Structures with X-ray and Neutron Diffraction Data from Single Crystals

BY ALEXANDER WLODAWER

National Measurement Laboratory, National Bureau of Standards, Washington, DC 20234, *USA*

AND WAYNE A. HENDRICKSON

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC 20375, *USA*

(*Received* 2 *July* 1981; *accepted* 6 *October* 1981)

Abstract

A procedure is presented for the stereochemically restrained least-squares refinement of macromolecular structures with neutron and X-ray diffraction data from single crystals. This procedure has been tested by refining a model of ribonuclease A using neutron data to minimal spacings of 2·8 Å and X-ray data from within 2·0 Å spacings. Joint X-ray and neutron refinement is well conditioned and tends to avoid false minima that may occur when a medium-resolution structure is refined solely with the neutron structure factors.

Introduction

Several methods for the refinement of the single-crystal neutron diffraction data collected on proteins have been tried in the last few years, but none of these has been completely satisfactory. The structures of metmyoglobin and carbonmonoxymyoglobin were refined by the real-space techniques at 2 Å (R = 32%) and 1·8 Å (R = 37%) resolution respectively (Schoenborn & Diamond, 1976; Norvell & Schoenborn, 1976). The structure of triclinic lysozyme was refined by Bentley & Mason (1981) using the least-squares technique of Agarwal (1978) and the idealization procedure of