Acta Cryst. (1980). A**36**, 657–669

Pseudoatom Expansions of the H_2 Electron Density

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

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

AND J. N. VARGHESE†

Crystallography Centre, The University of Western Australia, Nedlands, Western Australia 6009, *Australia*

(*Received* 1 *November* 1979; *accepted* 11 *February* 1980)

Abstract

Representations of a high-quality molecular electron density are studied. An evaluation of restricted radial functions is made using a least-squares figure of merit, the molecular quadrupole moment, the electric field at the nucleus, the electric-field gradient at the nucleus, an approximate energy and difference-density maps. It is concluded that for an accurate representation expansions to the quadrupole level or higher are needed with either single-term Slater-type radial functions having all exponents optimized, or two-term functions with a common exponent optimized. Generalized scattering factors from the restricted radial functions are not good approximations to those from completely flexible radial functions, but nevertheless give good property values.

Introduction

It is now a common practice (Coppens, 1977; Dawson, 1967; Fink, Gregory & Moore, 1976; Harel & Hirshfeld, 1975; Kohl & Bartell, 1969; Price, Varghese & Maslen, 1978; Stevens, 1979; Stewart, 1976; Vidal-Valat, Vidal & Kurki-Suonio, 1978) for molecular electron densities derived from X-ray and electron diffraction data to be expressed as a superposition of pseudoatom densities centred on each of the nuclei of the molecule. By denoting the nuclear position vectors as $\mathbf{R}_a$ and the associated pseudoatom density as ρ_a the total density becomes

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

The pseudoatom densities may be expressed as a finite multipole expansion about the nucleus (Dawson, 1965), with

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

where $\rho_{a,l}(r_a)$ is a radial density function, and $P_l^m(\cos\theta_a)$ is an associated Legendre polynomial. Stewart, Bentley & Goodman (1975) showed that such an expansion for diatomic molecules can be performed to arbitrary accuracy with the use of sufficiently flexible radial functions. The methods of that paper are readily used in producing accurate generalized scattering factors, which are the Fourier–Bessel transforms of the $\rho_{a,l}$'s in (2), from molecular wavefunctions, and have been applied to H_2, BF, CO, N_2 and the first-row diatomic hydrides.

As formulated by Stewart, Bentley & Goodman (1975), generalized scattering factors are equivalent to solving numerically for $\rho_{a,l}(r_a)$ (2) at each value of r_a and r_b. From such a development, Epstein, Bentley & Stewart (1977) have determined the importance of the various multipoles in an expansion, and have been able to make qualitative statements regarding the form of suitable radial basis functions for use in analysing X-ray and electron diffraction data. However, the generalized X-ray scattering factors just described cannot be obtained by fitting an experimental charge density with a pseudoatom expansion. Instead the experimentalist has to be content with a truncated set of multipoles and a severely limited set of radial functions, usually single- or two-term Slater-type functions, $r^n \exp(-\zeta_n r)$, or Gaussian-type functions $r^n \exp(-\alpha_n r^2)$. Hence we have preferred to deal directly with density functions using short expansions of Slater functions. In the present work various possibilities for such a procedure, many of which have been used in analysing experimental situations, will be critically examined in the context of the pseudoatom multipole analysis of a theoretical electron density for the hydrogen molecule. Some commonly used procedures give very poor expansions. It is the purpose of this paper to bring attention to these and to suggest better procedures. The number of Slater functions per multipole needed to fit the H_2 molecular density is investigated. The necessity of some form of exponent optimization is discussed along with the role of higher multipoles when less than completely flexible radial functions are employed.

* To whom all correspondence should be addressed.

† Present address: School of Molecular Sciences, University of Sussex, Falmer, BN1 9QJ, Sussex, England.

0567-7394/80/040657-13$01.00

Subsequent publications will extend the examination to the first-row hydrides and will apply the results to experimental situations.

Method

The calculated electron density for a diatomic molecule is represented as the sum of two nucleus-centred multipole expansions.

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

with P_j a Legendre polynomial and $\rho_{a,j}(r_a)$ a trial radial density function. By using Slater-type functions, we obtain

$$\rho_{a,j}(r_a) = \sum_{n=j}^{N_j} P_{n,j} N_{n,j} r_a^n \exp(-\zeta_n r_a). \quad (4)$$

Here $P_{n,j}$ is a variable population parameter, ζ_n, the exponent, may be optimized or held fixed at a value determined by some recipe, and n is an integer greater than or equal to the order of the Legendre polynomial with which it is associated. This constraint is imposed in order to ensure convergence of the integrals determining the electric field and field gradient, and to satisfy Poisson's equation for small r_a (Stewart, 1977). The normalizer, $N_{n,j}$, is that used by Hansen & Coppens (1978), with monopole terms normalized to unity, so that each monopole term contains $P_{n,0}$ electrons, and higher-order multipoles normalized so that their absolute value integrates to 2·0 electrons. This results in the populations $P_{n,j}$ having ready physical interpretations (Price, 1976). Thus the monopole populations directly give the number of electrons on each pseudoatom and higher multipole populations measure the quantity of charge which is redistributed.

After substituting (4) into (3) the density $\rho^c_{\rm mol}$ is fitted to the molecular electron density using the method of least squares. The least-squares error

$$\varepsilon = \int(\rho_{\rm mol} - \rho^c_{\rm mol})^2 \, d\tau \quad (5)$$

is minimized with respect to the population coefficients, $P_{n,j}$ of (4), and if required to the exponents ζ_n. Two quantitative measures of the goodness of fit are used. One is the least-squares figure of merit, R_w, defined as the square root of the relative error function

$$R_w = [\varepsilon/\int\rho^2_{\rm mol} \, d\tau]^{1/2}. \quad (6)$$

Since the calculated molecular density is not constrained to contain the number of electrons in the molecule, another measure is the fraction of electrons accounted for by the model, F, and given by the sum of the various monopole populations divided by the number of electrons in the molecule. Values of F significantly different from 1·0 indicate deficiencies in the pseudoatom expansion which may be in the radial or the angular sets of functions or in both. A renormalization of $\rho^c_{\rm mol}$ is made after the least-squares procedure by multiplying each population by $1/F$.

It is necessary to limit the values of n considered in (4) and also the number of higher multipoles considered. Equation (4) does not include what would seem to be a reasonable possibility; a radial expansion consisting of a linear combination of terms with equal values of n and hence differing only in their ζ_n. This is the more familiar situation met in basis-set expansions for atomic and molecular wavefunctions. Experiments with this type of density function resulted in least-squares refining to nearly equal exponents ζ_n. Further limitations on the n values considered have been guided by the fact that molecular electron densities are generated from Hartree–Fock calculations, which employ basis functions with low powers of r (Stewart, 1969). Hence we chose to consider expansions with different integral powers of r and the lowest powers consistent with the requirement that $n \geq l$. Examination of the multipoles was limited to those outlined in Table 1. Each multipole term in an expansion has the same number of radial functions. The radial functions investigated are shown in Table 2.

A number of ways of choosing the exponents ζ_n in (4) were studied and are described below.

(1) Scheme 1. $\zeta = 2{\cdot}0$. This corresponds to choosing the atomic hydrogen 1s exponent so that the exponent for the product $(1s)^2 = 2{\cdot}0$.

Table 1. *Multipole expansions investigated*

For H_2, $J = K$ by symmetry.

Expansion	Order of the highest Legendre polynomial $P_j(\cos\theta)$ in the expansion	Abbreviation used in the text for the expansion
Monopole terms only	0	M
M + dipole terms	1	D
D + quadrupole terms	2	Q
Q + octopole terms	3	O
O + hexadecapole terms	4	H

Table 2. *Radial expansions as in* (4) *for each multipole term in* (3)

Type of expansion	Powers of r for each multipole*
single term	M_0, D_1, Q_2, O_3, H_4
two term	$M_0M_1, D_1D_2, Q_2Q_3,$ O_3O_4, H_4H_5
three term	$M_0M_1M_2, D_1D_2D_3, Q_2Q_3Q_4,$ $O_3O_4O_5, H_4H_5H_6$

* The subscript to each multipole abbreviation is the power of r in the radial function.

(2) Scheme 2. $\zeta = 2{\cdot}48$. The zeta value here comes from the standard molecular exponent for hydrogen of Hehre, Stewart & Pople (1969).

(3) Scheme 3. All functions share a common exponent, ζ_{opt}, which is optimized.

(4) Scheme 4. All exponents are separately optimized.

These successive schemes for choosing exponents are very similar to those employed by Fraga & Ransil (1961) in a LCAO–SCF analysis of the H_2 molecule.

Exponent optimizations were achieved in simple cases by grid search methods and in the case of many variables with a gradient expansion algorithm (Bevington, 1969). The length of this procedure prevented us from fully studying three-term functions in scheme 4.

The electron density function, ρ_{mol}, for the $^1\Sigma_g^+$ ground state of H_2 is that of Stewart, Davidson & Simpson (1965), based on the first ten natural spin orbitals (Davidson & Jones, 1962) of the accurate wave-function of Kolos & Roothaan (1960). These ten natural spin orbitals account for 99% of the correlation energy. The density is given as

$$\rho_{mol} = 2\left(\frac{2}{R}\right)^3 \frac{1}{2\pi} \exp(-\alpha\xi) \sum_{k,j} a_{k,j}\, \xi^k\, \eta^j \qquad (7)$$

with $\alpha = 1{\cdot}99$, $R = 1{\cdot}4009$ a.u. Here ξ and η are confocal elliptical coordinates. The coefficients $a_{k,j}$ are tabulated by Stewart, Davidson & Simpson (1965). As written, (7) integrates to 1·999825 electrons and for the present work ρ_{mol} was rescaled to contain 2·0 electrons.

A comparison of physical properties calculated from ρ^c_{mol} with those obtained from ρ_{mol} is a measure of the effectiveness of any particular multipole expansion. Physical properties used for the comparison are the molecular quadrupole moment Q, the electric field at the nucleus ε_H, the electric-field gradient at the nucleus q_H, and the energy calculated with the approximate formula of Politzer (1979). Expressions for the above properties are given by

$$\begin{aligned} Q &= R^2 - \langle r^2 P_2(\cos\theta)\rangle, \\ \varepsilon_H &= -(1/R^2) + \langle r_a^{-2} P_1(\cos\theta_a)\rangle, \\ q_H &= (2/R^3) - 2\langle r_a^{-3} P_2(\cos\theta_a)\rangle, \end{aligned} \qquad (8)$$

where $\langle h\rangle = \int h\rho\, d\tau$.

The approximate energy is given by

$$E = \sum_A^{atoms} k_A Z_A V_A^{mol}, \qquad (9)$$

where V_A^{mol} is the electrostatic potential at the nucleus A, with charge Z_A. The parameter k_A comes from

$$k_A = E_A^{atom}/Z_A V_A^{atom} \qquad (10)$$

with E_A^{atom} chosen as the near-Hartree–Fock atomic energy. For hydrogen, k_H can be obtained exactly and is 0·5. Thus for H_2 (9) reduces to

$$E = V_H = (1/R) - \langle r_a^{-1}\rangle. \qquad (11)$$

One-centre integrals required in the evaluation of the properties (8) and (11) for ρ^c_{mol} are straightforward. Two-centre integrals were evaluated through an expansion used by Pitzer, Kern & Lipscomb (1962) of the solid spherical harmonic on one centre about the other centre. Accordingly, the two-centre integrals reduce to a sum of single-centre integrals which are readily determined. Care, however, must be taken with the electric-field gradient, as pointed out by Pitzer, Kern & Lipscomb (1962), to treat correctly singularities at $r_b = 0$.

Evaluation of molecular properties from the Stewart, Davidson & Simpson (1965) density function ρ_{mol} requires calculation of all integrals in confocal elliptical coordinates. The quadrupole moment is readily obtained and has been reported elsewhere (Stewart, Bentley & Goodman, 1975). However, the integrations for ε_H and q_H are more difficult, and to our knowledge have only been attempted in part before (Epstein, 1972). An outline of the evaluation is given in the Appendix, along with the evaluation of V_H.

Values of Q, ε_H, q_H and V_H from ρ_{mol} and the corresponding expectation values of the operators in (8) and (11) are given in Table 3. Both ε_H and q_H refer to the positive z direction chosen to point towards the other nucleus.

To supplement the examination of physical properties, difference-density plots are valuable for giving a qualitative indication of the efficiency of each multipole expansion and for showing how faithfully bonding distributions are being reproduced. Several molecular difference densities are in current use. We use the difference between ρ_{mol} and the superposition of the exact ground-state hydrogen-atom densities at the same internuclear distance, so that

$$\Delta\rho_{mol} = \rho_{mol} - \sum_{a,b} |\psi_{1s}^H|^2 \qquad (12)$$

with an analogous expression for $\Delta\rho^c_{mol}$. This definition ensures that plots indicate how each expansion represents the redistribution of the electron density accom-

Table 3. *Physical properties from the* H_2 *electron density* (7) *scaled to two electrons*

All values are in a.u.

Operator	Expectation value	Property	Value
$r^2 P_2(\cos\theta)$	0·52100	Q	1·44152
r_a^{-1}	1·82499	V_H	−1·11116
$r_a^{-2} P_1(\cos\theta_a)$	0·51113	ε_H	0·00158
$r_a^{-3} P_2(\cos\theta_a)$	0·19408	q_H	0·33930

panying bonding. Also, it is this difference density which is most commonly used in reporting X-ray and electron diffraction experiments. It is therefore of considerable interest to determine the number and type of angular and radial functions required to reproduce this feature for a simple molecule.

All calculations reported with exponent optimization have converged so that successive R_w values differ by less than 1×10^{-7}. This corresponds to changes of the order of 10^{-3} a.u. in the properties and the exponents.

Least-squares fits

Least-squares populations for one-term functions from schemes 1 to 4 are shown in Table 4. Corresponding data for two- and three-term functions are not given, as they are more difficult to interpret, and show behaviour similar to that of single-term functions. Results of exponent optimization for the single-term functions can be seen in Table 5, where all the exponents for schemes 3 and 4 differ significantly from either the 2·0 or 2·48 of schemes 1 and 2. Differences are greatest for the higher multipoles in scheme 4.

The populations in Table 4 are a direct measure of the relative importance of each of the multipole functions in each scheme. This is because single-term functions of the type employed in this analysis yield scattering factors whose peak height is independent of the exponent, but is proportional simply to the population. The constants of proportionality are 1·380, 1·096, 0·929 and 0·817 for the functions D_1, Q_2, O_3 and H_4 respectively.

Inspection of Table 4 shows that the relative importance of the multipoles depends on the choice of exponent for the radial function of each multipole. Even the sign of the multipole can alter. However, once the radial basis is made flexible enough by including exponent optimization in the manner of either schemes 3 or 4 then the importance of each multipole decreases with increasing order of the multipole. Thus octopole and hexadecapole terms become less important as the representation of the lower-order multipoles is improved.

In a generalized scattering-factor analysis of the same H_2 density employed in this paper, Epstein, Bentley & Stewart (1977) concluded that the octopole scattering factor is too small to be of significance in the analysis of electron diffraction data. However, the data of Tables 4 and 5 emphasize that this is only likely to be true provided some form of exponent optimization is performed to give the lower-order multipoles sufficiently flexible radial functions.

There are three factors affecting the least-squares goodness of fit and the accuracy of the calculated physical properties. They are the scheme used for determining exponents, the number of terms in each radial expansion, and the length of the multipole

Table 5. *Optimized exponents of single-term radial functions in schemes* 3 *and* 4

For an explanation of the notation see Tables 1 and 2.

(*a*) Scheme 3

	M	*D*	*Q*	*O*	*H*
ζ	2·252	2·275	2·231	2·219	2·211

(*b*) Scheme 4

	M	*D*	*Q*	*O*	*H*
M_0	2·252	2·243	2·219	2·200	2·195
D_1		2·961	2·612	2·404	2·378
Q_2			3·098	2·789	2·749
O_3				3·537	3·381
H_4					4·259

Table 4. *Least-squares populations for single-term radial functions in schemes* 1 *to* 4

For an explanation of the notation see Tables 1 and 2.

Scheme 1

	M	*D*	*Q*	*O*	*H*
M_0	1·000	1·000	1·000	1·000	1·000
D_1		0·157	0·216	0·235	0·262
Q_2			0·143	0·136	0·105
O_3				0·079	0·158
H_4					0·138

Scheme 2

	M	*D*	*Q*	*O*	*H*
M_0	1·000	1·000	1·000	1·000	1·000
D_1		0·140	0·147	0·143	0·132
Q_2			0·032	0·033	0·046
O_3				−0·016	−0·038
H_4					−0·071

Scheme 3

	M	*D*	*Q*	*O*	*H*
M_0	1·000	1·000	1·000	1·000	1·000
D_1		0·143	0·168	0·176	0·182
Q_2			0·077	0·078	0·074
O_3				0·028	0·042
H_4					0·027

Scheme 4

	M	*D*	*Q*	*O*	*H*
M_0	1·000	1·000	1·000	1·000	1·000
D_1		0·088	0·112	0·138	0·142
Q_2			0·031	0·049	0·051
O_3				0·013	0·015
H_4					0·003

expansion, Table 1. As the easiest way to appreciate trends with these factors is graphically, Figs. 1 to 6 show respectively R_w, F, Q, ε_H, q_H and E calculated from Politzer's (1979) approximate relation, graphed against the length of the multipole expansion, for each of schemes 1 to 4 and for each of the types of radial expansions used.

In discussing R_w (Fig. 1) we make the arbitrary judgement that a good fit will have $R_w \leq 0{\cdot}02$. With this in mind it can be clearly seen that the use of single-term radial functions is inadequate in schemes 1, 2 or 3 and full optimization of exponents is required for such functions to be satisfactory, Fig. 1(*d*). When all exponents are optimized there is little difference between one- and two-term functions and nothing is gained from the labour involved in handling the much larger set of parameters for the latter set. Without full optimization there is a considerable difference between one-term functions and the other sets (see Fig. 1*a–c*), but both two- and three-term functions give similar R_w values and very satisfactory fits to the data. Thus the major deficiency of one-term functions can largely be accounted for by adding one extra function with the same exponent, but with the next-higher power of r. All graphs of R_w show that a multipole expansion up to the quadrupole level is essential to obtain $R_w \leq 0{\cdot}02$. The graphs of F (Fig. 2) amplify these conclusions.

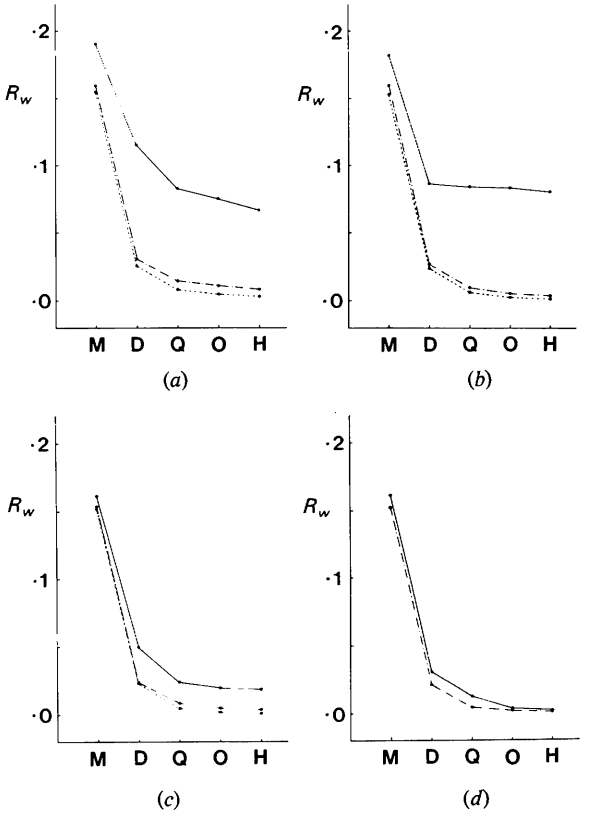

Fig. 1. The least-squares figure of merit R_w plotted against multipole expansion length. One-term functions (——), two-term functions (— - — - —), three-term functions (– – –); (*a*) scheme 1, (*b*) scheme 2, (*c*) scheme 3 and (*d*) scheme 4. For an explanation of the notation on the abscissa see Table 1.

Physical properties

Low R_w values are not sufficient to guarantee that a given physical property will be satisfactorily reproduced by a pseudoatom expansion. Hence it is instructive to consider some physical properties, $\langle h \rangle = \int h \rho^c_{\text{mol}} \, d\tau$, calculated from the rescaled densities ρ^c_{mol}. Since the ability of the pseudoatom expansions to reproduce accurately the electron density is under examination, it is appropriate to compare the representation of different physical properties with each expansion in terms of a range in the respective electronic expectation values from ρ_{mol}. Using D, the percentage difference, between the respective operator expectation value calculated for ρ_{mol} and ρ^c_{mol}, so that, for example, with the quadrupole moment

$$D = \frac{[\langle r^2 P_2(\cos\theta)\rangle_{\rho_{\text{mol}}} - \langle r^2 P_2(\cos\theta)\rangle_{\rho^c_{\text{mol}}}]}{\langle r^2 P_2(\cos\theta)\rangle_{\rho_{\text{mol}}}} \times 100, \tag{13}$$

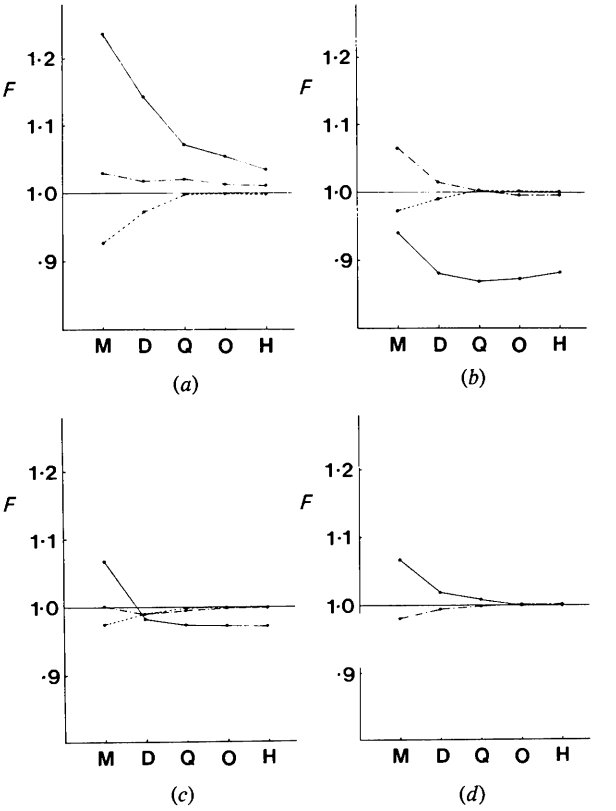

Fig. 2. The fraction of electrons counted, F. For an explanation of the representation and notation see Fig. 1.

we obtain a direct measure of the adequacy of the contribution of the calculated density to the property. Considering a range of $\pm 10\%$ in D to be of satisfactory accuracy for these expectation values, then this corresponds to a range of $\pm 0{\cdot}18$ a.u. in E, $\pm 0{\cdot}05$ a.u. in Q and ε_H and $\pm 0{\cdot}04$ a.u. in q_H.

Graphs of the molecular quadrupole moment, Q, are displayed in Fig. 3. The quadrupole moment is inadequately represented in scheme 1 and is only well reproduced in schemes 2 and 3 for two- and three-term functions in multipole expansions containing quadrupole or higher-order terms. With all exponents fully optimized (Fig. 3*d*), Q is well reproduced by single-term functions at the dipole level. Using two-term functions and larger multipole expansions provides little improvement.

The electric field at the nucleus, ε_H, which must be zero for the equilibrium internuclear distance and the exact density function has a value from ρ_{mol} close to this. Only single-term functions in schemes 1 and 2, and expressions consisting of monopoles alone fail to yield satisfactory values (see Fig. 4). When some exponent optimization is included all expansions above the monopole level give very close agreement with the theoretical value. Once again the similarity between two- and three-term functions for schemes 1 to 3 and one- and two-term functions in scheme 4 is evident.

New features appear with q_H, shown in Fig. 5. It has not been possible to represent q_H adequately with a multipole expansion below the quadrupole level, as expected from the general considerations of Stewart, Bentley & Goodman (1975). Above the quadrupole level one-term functions provide a reasonable representation of q_H in schemes 2 and 3 and give close to the accurate value in 4. The similarities between the functions noted already are repeated.

Values of the approximate energy, $E = V_H$, displayed in Fig. 6, show similar behaviour to F. Furthermore, it can be seen on inspection that the product FE is invariably close to E obtained from ρ_{mol}, with an average of $-1{\cdot}112$ for all density basis sets considered.

Difference-density maps

The trends already discussed are also shown by the difference-density plots $\Delta\rho^c_{mol}$ which are reproduced in Figs. 8 to 11 for all the schemes studied. In addition, these plots provide some new insights of their own.

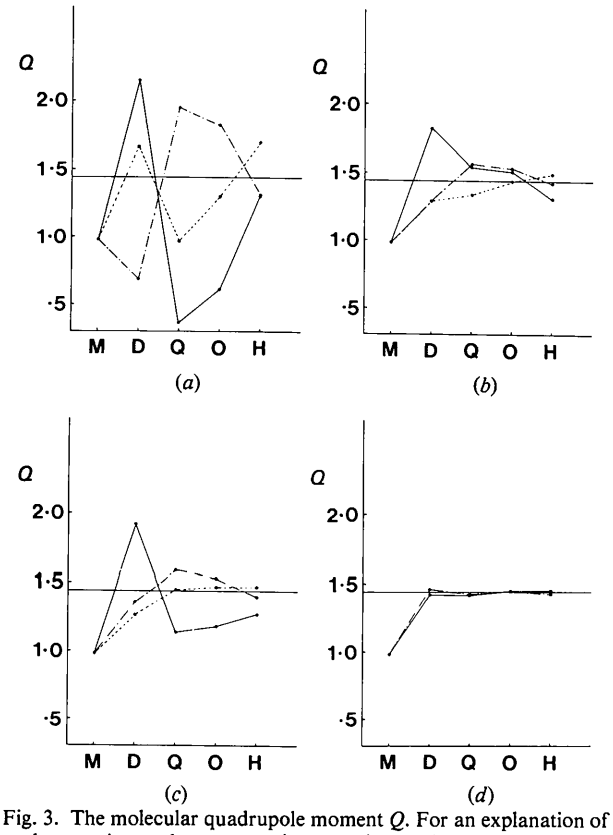

Fig. 3. The molecular quadrupole moment Q. For an explanation of the notation and representation see Fig. 1. The horizontal line is Q from ρ_{mol} (7). Q is in a.u.

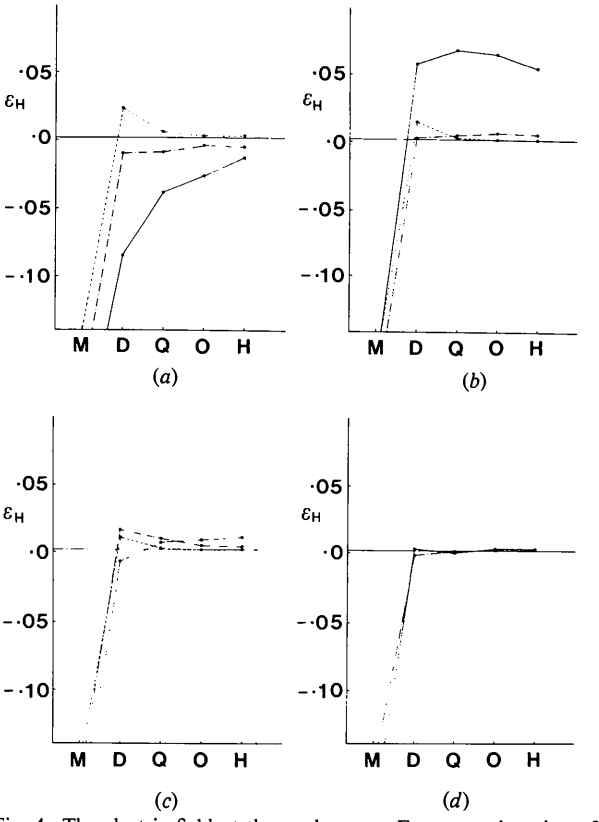

Fig. 4. The electric field at the nucleus, ε_H. For an explanation of the notation and representation see Fig. 1. The horizontal line is ε_H from ρ_{mol} (7). ε_H is in a.u.

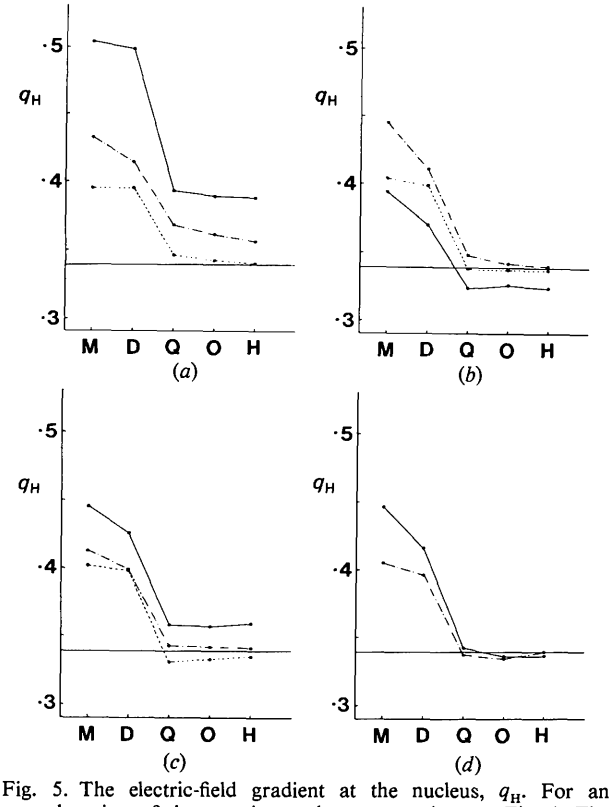

Fig. 5. The electric-field gradient at the nucleus, q_H. For an explanation of the notation and representation see Fig. 1. The horizontal line is q_H from ρ_{mol} (7). q_H is in a.u.

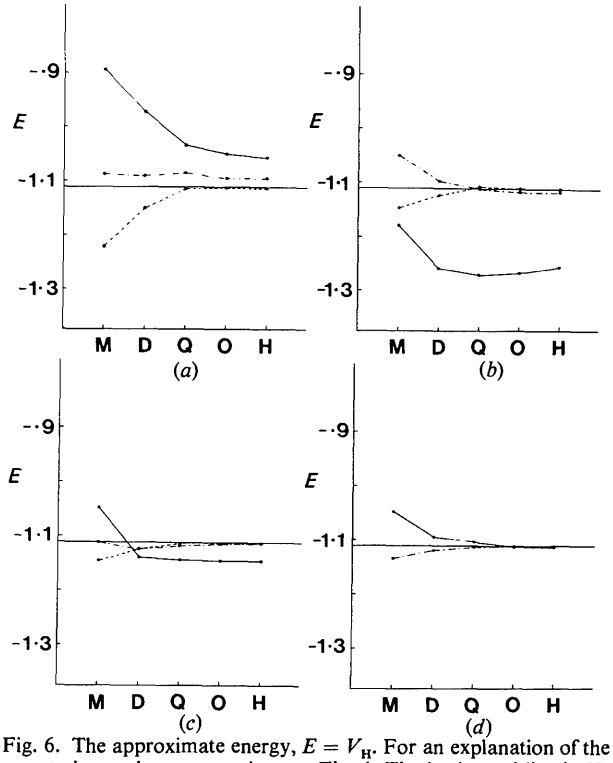

Fig. 6. The approximate energy, $E = V_H$. For an explanation of the notation and representation see Fig. 1. The horizontal line is V_H from ρ_{mol} (7). E is in a.u.

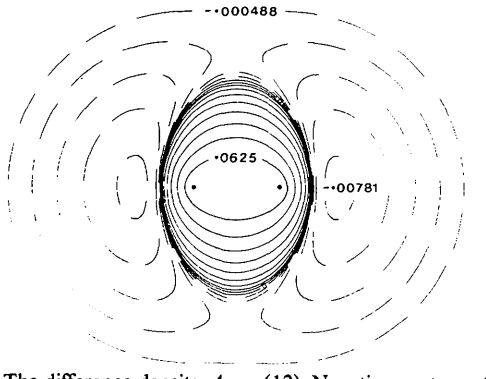

Fig. 7. The difference density $\Delta\rho_{mol}$ (12). Negative contours (---), positive contours (——). Successive contours differ by a factor of two. The densities marked are given in e a_0^{-3} (1 e a_0^{-3} = 6·74834 e Å^{-3}).

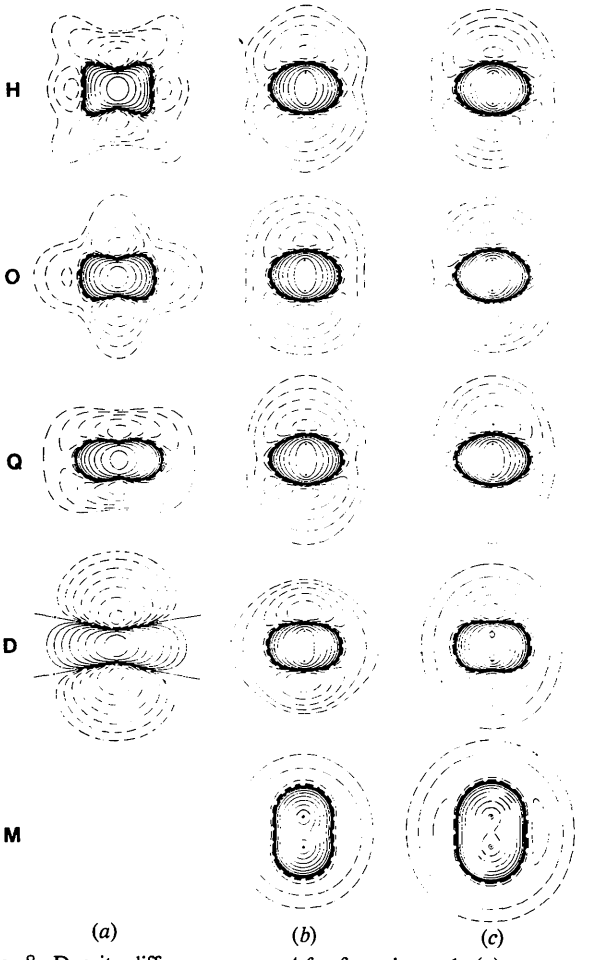

Fig. 8. Density-difference maps $\Delta\rho^c_{mol}$ for scheme 1, (*a*) one-term functions, (*b*) two-term functions, (*c*) three-term functions; heavy dots show the nuclei with the internuclear axis vertical. For an explanation of the notation *M*, *D*, *etc.* see Table 1. Outer negative contours are the same as in Fig. 7 except (*a*) *D* where it is −0·000244 e a_0^{-3}. The highest positive contours in (*c*) *M* and (*c*) *D* (*i.e.* the very small inner contours) represent 0·125 e a_0^{-3} and that in (*a*) *D* represents 0·0313 e a_0^{-3}, the remainder are the same as Fig. 7.

The accurate difference density is shown in Fig. 7. It can be seen to exhibit familiar bonding features with a large accumulation of electron density around the bond and immediately behind the nuclei. A corresponding decrease of electron density occurs beyond this region, with depressions occurring some distance behind the nuclei. These depressions are near 0·01 e a_0^{-3} deep and the build up of density in the bond peaks is approximately 0·1 e a_0^{-3}. Reproduction of the general features of $\Delta\rho_{\rm mol}$ places no more stringent demands on the pseudoatom expansion than the requirements already discussed. And so it can be seen (Figs. 8 to 10) that one-term functions in any of schemes 1 to 3 fail to reproduce the features of the $\Delta\rho_{\rm mol}$ plot in magnitude or shape, but if the multipole expansion is either at the quadrupole level or better, then two-term functions satisfactorily reproduce the reference in all schemes. Nevertheless, there are still noticeable differences between the better $\Delta\rho^c_{\rm mol}$ plots and $\Delta\rho_{\rm mol}$, principally in the shape of the outer contours and in the depth of the depression behind the nucleus. Three-term functions are required to account accurately for these features if all exponents are not fully optimized, Fig. 11.

Comparison with generalized scattering factors

As mentioned in the *Introduction*, the generalized scattering factors (GSF's), $f_l^{\rm GSF}(S)$, of Stewart, Bentley & Goodman (1975) have been recommended as a guide to the rational choice of pseudoatom multipoles and radial basis functions (Epstein, Bentley & Stewart, 1977). In essence it is expected that the structure of each $f_l^{\rm GSF}(S)$ curve can suggest the number of exponential-type radial functions for each multipole, and the relative magnitude of the $f_l^{\rm GSF}(S)$ can be used to decide on the priority for introduction of the multipoles into a least-squares analysis of electron diffraction for diatomic molecules.

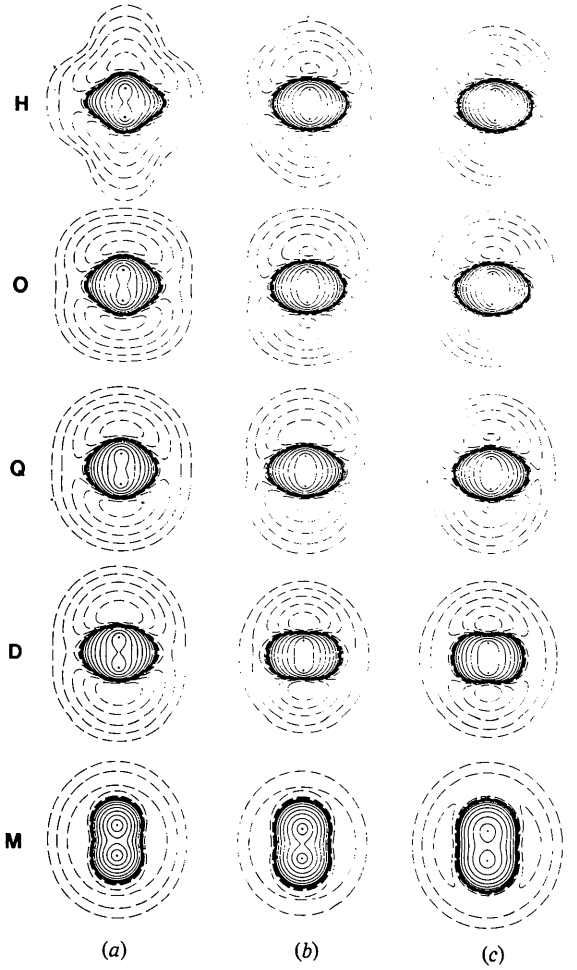

Fig. 9. Density-difference maps $\Delta\rho^c_{\rm mol}$ for scheme 2. The highest positive contours for all (*a*) represent 0·125 e a_0^{-3}.

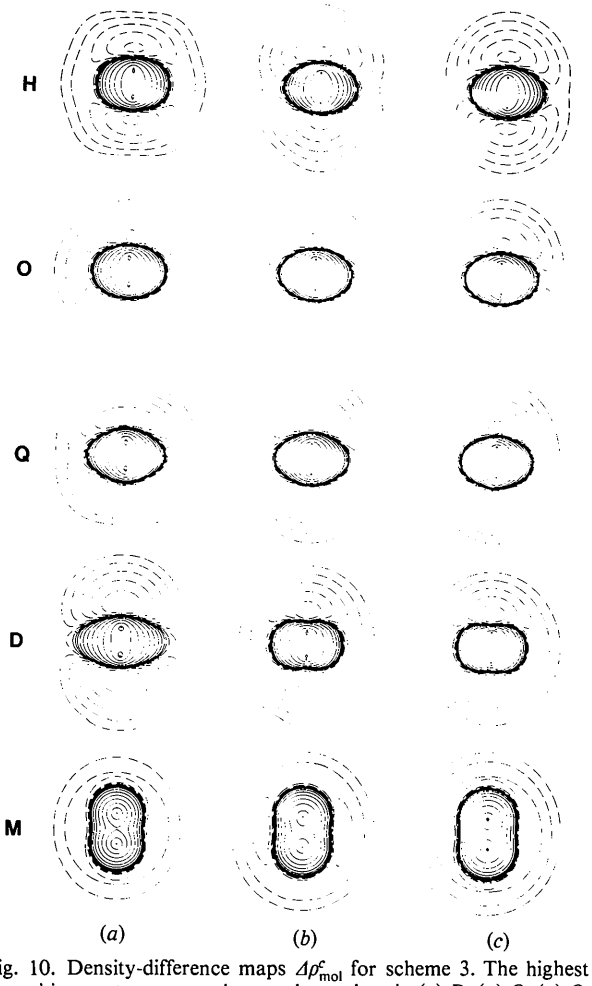

Fig. 10. Density-difference maps $\Delta\rho^c_{\rm mol}$ for scheme 3. The highest positive contours very close to the nucleus in (*a*) *D*, (*a*) *Q*, (*a*) *O*, and (*a*) *H* represent 0·125 e a_0^{-3}.

It is therefore instructive to compare the results of the foregoing section with GSF's obtained from the same density function for H_2. Since GSF's are calculated for a finite range of S, any inverse Fourier transformation to the corresponding real-space radial functions will include series-termination effects. The resulting multipole in real space will be the convolution of the true $\rho_{a,l}$ with a first-order spherical Bessel function. To avoid this, GSF's are compared with sets of restricted radial functions in reciprocal space.

Generalized scattering factors for H_2 up to the quadrupole level are tabulated by Stewart, Bentley & Goodman (1975). The calculated scattering factor for the lth multipole on centre a is given by

$$f^c_{a,l}(s) = 4\pi \int_0^\infty \rho_{a,l}(r_a)\, j_l(sr_a) r_a^2 \, dr_a. \qquad (14)$$

With an expansion of $\rho_{a,l}(r_a)$ in k terms of the form

$$\rho_{a,l}(r_a) = \sum_{i=l}^{l+k-1} P_i N_i r_a^i \exp(-\zeta_i r_a), \qquad (15)$$

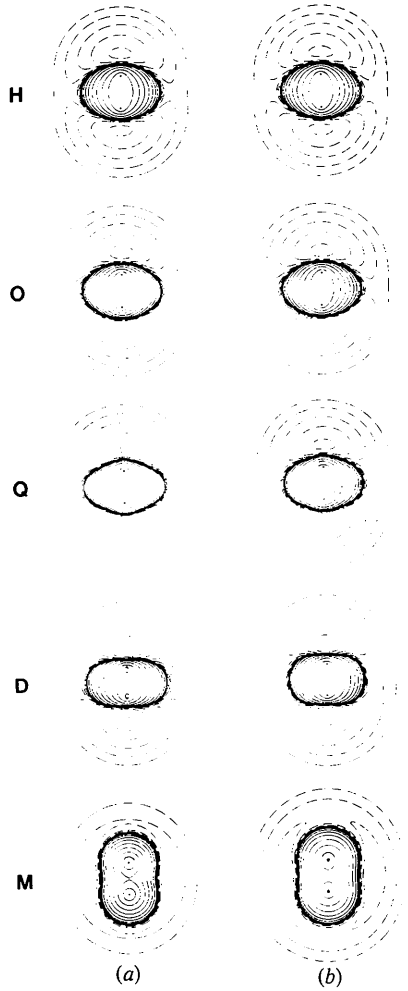

Fig. 11. Density-difference maps $\Delta\rho^c_{\rm mol}$ for scheme 4. The highest positive contour in (*a*) *D* represents 0·125 e a_0^{-3}.

(14) can be written

$$f^c_{a,l}(s) = 4\pi \sum_{i=l}^{l+k-1} P_i N_i J_{i+2,l}(s,\zeta_i) \qquad (16)$$

with

$$J_{n,m}(s,\alpha) = \int_0^\infty r^n \exp(-\alpha r) j_m(sr)\, dr. \qquad (17)$$

The $J_{n,m}$ can be obtained from the recurrence relations published by Avery & Watson (1977), beginning with $J_{1,0} = (s^2 + \alpha^2)^{-1}$.

It is readily established that the GSF's themselves can be fitted from a small expansion of exponential radial functions. The adequacy of a fit to a GSF curve of order l is measured by the relation

$$R_l = \left\{ \chi_l^2 / \sum_{s_i} [f_l^{\rm GSF}(s_i)]^2 \right\}^{1/2}, \qquad (18)$$

where

$$\chi_l^2 = \sum_{s_i} [f_l^{\rm GSF}(s_i) - f_l^c(s_i)]^2 \qquad (19)$$

with

$$f_l^c(s_i) = \sum_{j=l}^{l+k-1} c_j J_{j+2,l}(s_i, \zeta_j) \qquad (20)$$

and the summation in s_i is over the set of sin θ/λ values tabulated by Stewart, Bentley & Goodman (1975). Least-squares parameters obtained from this analysis are given in Table 6. The P_i are populations of normalized radial functions which have been obtained

Table 6. *Least-squares parameters and physical properties obtained for fits to GSF's in* H_2 *up to the quadrupole level*

The population coefficients, P_i, are given in brackets.

	One term	Two term	Three term	Multipole
ζ_i	2·265 (1·001)	2·472 (0·737)	2·503 (0·654)	M_0
		2·492 (0·262)	2·485 (0·509)	M_1
			2·944 (−0·223)	M_2
ζ_i	2·774 (0·097)	2·350 (0·203)	2·530 (0·063)	D_1
		2·610 (−0·105)	3·370 (0·124)	D_2
			3·832 (−0·090)	D_3
ζ_i	3·745 (0·014)	3·014 (0·059)	3·324 (0·005)	Q_2
		3·464 (−0·044)	4·117 (0·054)	Q_3
			4·726 (−0·044)	Q_4
R_w	0·0206	0·0120	0·0051	
F	1·001	0·999	1·000	
Q	1·441	1·437	1·444	
ε_H	0·004	0·009	−0·001	
q_H	0·344	0·331	0·338	
R_0	0·0420	0·0227	0·0022	
R_1	0·0150	0·0072	0·0015	
R_2	0·0833	0·0267	0·0149	

from the c_i of (20) so that direct comparison can be made with Tables 4 and 7. The R_0 factor quoted for the monopole term is for $\Delta f_0^{GSF}(s)$, the difference between $f_0(s)$ and the scattering factor of an exact ground-state hydrogen atom replacing $f_0^{GSF}(s)$ in (19), although the actual least-squares fitting was made on the $f_0^{GSF}(s)$ curve. Table 6 also gives the values of physical properties calculated from the radial functions fitted to the GSF's. Excellent agreement with the theoretical property values is obtained for all the functions tried. Both two- and three-term functions show good fits to the individual multipole GSF's as seen from the small R values. However, one-term functions give poor fits. A more detailed examination of the one-term functions reveals that in the case of the monopole GSF the major disagreement occurs at large values of sin θ/λ, but for the quadrupole GSF there are differences over most of the range of sin θ/λ values.

Table 7. *Least-squares parameters and physical properties obtained for two- and three-term functions fitted directly to ρ_{mol} for H_2 up to the quadrupole level*

The P_i values are given in brackets.

	Two term	Three term	Multipole
ζ_i	2·285 (0·884)	2·249 (0·933)	M_0
	3·110 (0·116)	3·803 (0·075)	M_1
		8·324 (−0·008)	M_2
ζ_i	2·647 (0·065)	2·679 (0·091)	D_1
	4·374 (0·027)	3·924 (−0·016)	D_2
		5·042 (0·020)	D_3
ζ_i	4·574 (0·012)	2·666 (0·075)	Q_2
	10·581 (−0·001)	2·489 (−0·023)	Q_3
		3·911 (−0·039)	Q_4
R_w	0·0041	0·0033	
F	0·998	1·000	
Q	1·426	1·456	
ε_H	0·001	0·001	
q_H	0·337	0·332	

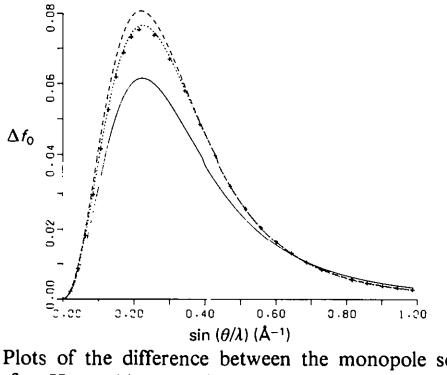

Fig. 12. Plots of the difference between the monopole scattering factor for H_2, $f_0(s)$, and the scattering factor of an exact ground-state hydrogen atom. One-term function (——), two-term function (· · · · ·), three-term function (– – –), results for the generalized scattering factors are given by crosses +.

A comparison of the parameters in Table 6 with those in Tables 4, 5 and 7 reveals that there are considerable differences between the radial functions obtained by fitting the molecular electron density with schemes 3 and 4 and those fitted to the GSF's. For the sake of comparison, parameters for a three-term function fitted to the molecular density are given in Table 7. These parameters have not converged to the same limits of R_w which were used in obtaining all other restricted radial functions. In terms of the parameter values even the three-term functions appear very different, but this is more apparent than real as shown in Figs. 12 to 14 which give plots of GSF's and scattering factors from the restricted functions.

The monopole scattering factors are presented as $\Delta f_0(s)$ in Fig. 12. Only the three-term function is a close approximation to $\Delta f_0^{GSF}(s)$, at low sin θ/λ, but the two- and three-term functions are both very good at high sin θ/λ. The dipole functions shown in Fig. 13 exhibit similar behaviour. The quadrupole term in Fig. 14, however, is different. None of the restricted radial functions is a good approximation to $f_2^{GSF}(s)$, but they do approach the GSF as the number of terms in the radial expansion increases. The single-term function is more than twice the value of $f_2^{GSF}(s)$ at low sin θ/λ, but is too small at sin θ/λ above 0·7 Å^{-1}. The two-term

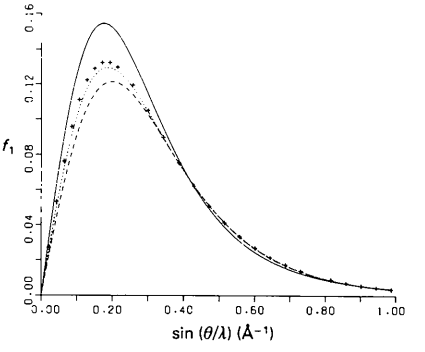

Fig. 13. Dipole scattering factors for H_2. The representation is the same as in Fig. 12.

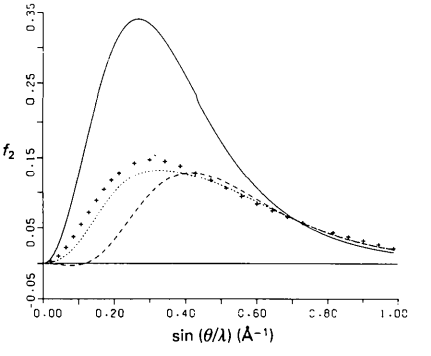

Fig. 14. Quadrupole scattering factors for H_2. The representation is the same as in Fig. 12.

$f_2^c(s)$ is negative below 0·12 Å^{-1}, and the three-term $f_2^c(s)$ is the only expansion to represent approximately $f_2^{GSF}(s)$ at all sin θ/λ values. Even this expansion is quite deficient at low sin θ/λ.

These results suggest that it is not possible to get a reasonable approximation to GSF's in an analysis of the molecular density with single-term exponential-type functions, although single-term functions are adequate for directly fitting GSF's. Generalized scattering factors are the multipole scattering factors and hence radial functions in real space, expected from a least-squares analysis of diffraction data in the limit of totally flexible radial basis sets. Unfortunately they are not close to what is obtained from analyses of electron densities, and hence diffraction data, with restricted sets of radial functions. Hence the usefulness of GSF's for determining the suitability of different sets of single- or several-term exponential functions or the magnitude of the various multipoles when using limited radial functions would seem to be somewhat restricted. Flexibility in the radial functions is the key to reproducing GSF's and as it increases the agreement with the GSF's rapidly improves. With two- and three-term functions the high sin θ/λ regions are well approximated. In real space this corresponds to a good fit to ρ_{mol} near the nuclei and in the region of the bond, but poor agreement at large distances from the bond, where even more flexible radial functions would seem to be required. This is undoubtedly why the inner moments of the electron density which appear in ε_H, q_H and V_H, are well represented by the limited radial-function fits (see Table 7 and Figs. 4 to 6) but Q which depends more heavily on the outer regions of the electron density is not so well reproduced. As Figs. 12 to 14 show, fitting directly to the density with restricted radial functions fits the high sin θ/λ region adequately, but not in the region below 0·5 Å^{-1}, and thus yields poorer values of Q (Table 7). On the other hand, similar fits to the GSF curve directly fit better at low sin θ/λ and therefore yield a good representation of Q, as seen from Table 6.

Conclusions

The results obtained in this work are only for the hydrogen molecule, a simple molecule with no core electrons. However, any inferences drawn will certainly represent the minimum requirements for nuclear-centred multipole expansions of more complicated molecules. Certain definite conclusions come from the results discussed.

Firstly, under most circumstances it is necessary to include at least up to quadrupole terms in a multipole expansion. The exceptions are for a description of the electric field at the nucleus and the molecular quadrupole moment in the scheme where all exponents are optimized, so that for practical purposes it would be unwise to work with less than quadrupole terms. Even though satisfactory for properties, the expansions to the quadrupole level still show discernible distortions in the $\Delta\rho_{mol}^c$ plots (see Figs. 8 to 11) although they only represent shifts in very small fractions of an electronic charge. It is necessary to include octopole terms to rectify this. Secondly, it is readily seen from the graphs and plots that one-term Slater radial functions are suitable only if all exponents are separately optimized. At this level it is certainly not necessary to use two-term Slater radial functions. If full optimization is not performed it is still possible to get a good fit to the density, but two-term Slater radial functions must be used. It is necessary then to use at least standard molecular exponents but it is preferable to optimize the common exponents.

Our results may be compared with those of Epstein, Bentley & Stewart (1977) who, working with generalized scattering factors obtained at the quadrupole and octopole levels for H_2, noted that a single exponential function should suffice to represent the generalized scattering factors for each multipole. This is at variance with our findings which clearly show that the GSF's cannot be reproduced by a single exponential fit to the electron density. The H_2 electron density itself, though, is successfully fitted with such a function, but it is emphasized that all exponents must be optimized. Epstein, Bentley & Stewart (1977) also observe that the octopole term appears to be too small to be of significance in the analysis of real diffraction data. Our results suggest that although this may be the case for totally flexible radial functions, radial functions with restricted flexibility yield appreciable coefficients for such high-order multipoles. Moreover, any difference densities obtained from a truncated multipole expansion will show features which are a direct result of the neglect of octopole terms.

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

We would like to thank Dr E. N. Maslen for suggesting the problem to us.

We wish to thank the Australian Grants Commission for support.

APPENDIX

A calculation of V_H, ε_H and q_H from ρ_{mol} (7) requires determination of the expectation values $\langle 1/r_a\rangle$, $\langle r_a^{-2} \times P_1(\cos\theta_a)\rangle$ and $\langle r_a^{-3} P_2(\cos\theta_a)\rangle$. V_H is obtained from

$$\langle 1/r_a\rangle = (2/R)\langle 1/(\xi+\eta)\rangle$$

$$= \frac{4}{R}\sum_{k,j} a_{kj}\left[\int_1^\infty \exp(-\alpha\xi)\,\xi^{k+1}\,d\xi\right.$$

$$\times \int_{-1}^{1} \eta^j \, d\eta - \int_{1}^{\infty} \exp(-\alpha\xi)\, \xi^k \, d\xi \int_{-1}^{1} \eta^{j+1} \, d\eta \Bigg], \tag{A1}$$

but since j is always even,

$$\int_{-1}^{1} \eta^{j+1} \, d\eta = 0,$$

hence $$\langle 1/r_a \rangle = (8/R) \sum_{k,j} (a_{kj}/j + 1) A_{k+1}(\alpha) \tag{A2}$$

with $$A_n(\alpha) = \int_{1}^{\infty} t^n \exp(-\alpha t)\, dt \tag{A3}$$

and $k = 0, 1, 2, \ldots, 12, j = 0, 2, 4, \ldots, 12$.

The determination of ε_H requires the evaluation of

$$\langle r_a^{-2} P_1(\cos\theta_a) \rangle = \int r_a^{-2} P_1(\cos\theta_a)\, \rho_{\text{mol}}\, d\tau. \tag{A4}$$

Writing

$$\frac{\cos\theta_a \, d\tau}{r_a^2} = \left(\frac{\xi\eta + 1}{\xi + \eta}\right)\left(\frac{2}{R}\right)^2 \frac{1}{(\xi+\eta)^2} \times \left(\frac{R}{2}\right)^3 (\xi^2 - \eta^2)\, d\xi\, d\eta\, d\varphi$$

$$= (R/2)(\xi\eta + 1)(\xi - \eta)\, d\xi\, d\eta\, d\varphi/(\xi+\eta)^2 \tag{A5}$$

and substituting (A5) and ρ_{mol} (7) into (A4), we get

$$\langle r_a^{-2} P_1(\cos\theta_a) \rangle = 2(2/R)^2 \sum_{k,j} a_{kj} \int_{1}^{\infty} d\xi \int_{-1}^{1} d\eta \exp(-\alpha\xi) \times [1/(\xi+\eta)^2][\xi^{k+2}\eta^{j+1} + \xi^{k+1}\eta^j - \xi^{k+1}\eta^{j+2} - \xi^k \eta^{j+1}]. \tag{A6}$$

Similarly for q_H

$$\langle r_a^{-3} P_2(\cos\theta_a) \rangle = (2/R)^3 \sum_{k,j} a_{kj} \times \int_{1}^{\infty} d\xi \int_{-1}^{1} d\eta \exp(-\alpha\xi)[1/(\xi+\eta)^4][3\xi^{k+3}\eta^{j+2} - 3\xi^{k+2}\eta^{j+3} + 5\xi^{k+2}\eta^{j+1} - 5\xi^{k+1}\eta^{j+2} + 3\xi^{k+1}\eta^j - 3\xi^k\eta^{j+1} - \xi^{k+3}\eta^j + \xi^k\eta^{j+3}]. \tag{A7}$$

Both integrals have singularities at nucleus a, $\eta = -1$, $\xi = 1$. In the evaluation it is necessary to integrate outside a sphere of radius ε on nucleus a and then to investigate the limit as $\varepsilon \to 0$, so that the integration is taken over the following limits (Isiguro, 1948; Reid & Vaida, 1973):

$$\lim_{\varepsilon\to 0}\left[\int_{1+\varepsilon}^{\infty} d\xi \int_{-1}^{1} d\eta + \int_{1}^{1+\varepsilon} d\xi \int_{\varepsilon-\xi}^{1} d\eta\right]. \tag{A8}$$

The first term in (A8) is a straightforward integration for the integrands of (A4) and (A7) yielding the same result after taking the limit as evaluating

$$\int_{1}^{\infty} d\xi \int_{-1}^{1} d\eta$$

directly. After considerable manipulation this contribution to ε_H is found to be

$$2(2/R)^2 \sum_{k,j} a_{kj} \Bigg\{(2j+3) L_{j+k+2}(\alpha) - (2j+1) L_{j+k}(\alpha) - 2(j+3) A_{j+k+1}(\alpha) + \sum_{l=1}^{j-1}{}' \left[\left(\frac{4l+2}{j-l} - \frac{2l}{j-l+2}\right) A_{k+l}(\alpha) - 2\frac{(l+1)}{j-l} A_{k+l+2}(\alpha)\right]\Bigg\} \tag{A9}$$

with $A_n(\alpha)$ given by (A3) and

$$L_m(\alpha) = \int_{1}^{\infty} t^m \exp(-\alpha t) \ln\left(\frac{t+1}{t-1}\right) dt. \tag{A10}$$

The primed summation $\sum'$ indicates a sum over odd values of l only, and thus only contributes when $j \geq 2$. The same two integrals $A_n(\alpha)$ and $L_m(\alpha)$ arise in a determination of the electric-field gradient from a simple H_2 wavefunction (Isiguro, 1948).

Similarly, the contribution to q_H is

$$(2/R)^3 \sum_{k,j} a_{kj} G_{kj}(\alpha), \tag{A11}$$

where

$$G_{kj}(\alpha) = [-(j+2)(j+1)(2j+3)/2] L_{j+k+2}(\alpha) + (2j+1)(j^2+j+1) L_{j+k}(\alpha) - [j(j-1)(2j-1)/2] L_{j+k-2}(\alpha) + (j+1)(2j^2+7j+6) A_{j+k+1}(\alpha) - j(2j^2+7j-1) A_{j+k-1}(\alpha) + \sum_{m=3}^{j-1}{}' \frac{2m(m-1)}{(j-m)(j-m+2)} [(m+j-1) \times A_{k+m-2}(\alpha) - (m+j+1) A_{k+m}(\alpha)]. \tag{A12}$$

Here $A_n(\alpha)$ and $L_m(\alpha)$ are given in (A3) and (A10) and again the primed summation includes only odd values of m. The $A_n(\alpha)$ are readily obtainable from tables of definite integrals. For the logarithmic integral $L_m(\alpha)$ we follow Epstein (1972) and split it into two parts, so that

$$L_m(\alpha) = \int_1^\infty \exp(-\alpha t)\, t^m \ln(t+1)\, dt - \int_1^\infty \exp(-\alpha t)\, t^m \ln(t-1)\, dt = LA_m(\alpha) - LB_m(\alpha). \quad (A13)$$

Then it is readily shown that

$$LA_m(\alpha) = A_m(\alpha) \ln 2 + \sum_{k=0}^{m} \frac{m! \exp(-\alpha)}{\alpha^{m+1}(m-k)} \times \left[\sum_{n=0}^{m-k-1} \frac{(2\alpha)^n}{n!} \sum_{l=0}^{k} \frac{(-\alpha)^l}{l!} \right] \quad (A14)$$

and

$$LB_m(\alpha) = -(C + \ln \alpha) A_m(\alpha) + \frac{m!}{\alpha^m} \sum_{k=0}^{m-1} \frac{A_k(\alpha)\, \alpha^k}{k!(m-k)} \quad (A15)$$

where C is Euler's constant.

It is now necessary to examine the behaviour of the second term in (A8). For $\langle r_a^{-2} P_1(\cos\theta_a)\rangle$ the limit as $\varepsilon \to 0$ vanishes. Thus, $\langle r_a^{-2} P_1(\cos\theta_a)\rangle$ is given by (A9). In the case of the electric-field gradient though, the limit leaves an extra term of $\frac{2}{3}\exp(-\alpha)$ which when included with (A11) gives,

$$\langle r_a^{-3} P_2(\cos\theta_a)\rangle = (2/R)^3 \sum_{k,j} a_{kj} G_{kj}(\alpha) + \tfrac{2}{3}\exp(-\alpha), \quad (A16)$$

which can be rewritten as

$$\langle r_a^{-3} P_2(\cos\theta_a)\rangle = (2/R)^3 \sum_{k,j} a_{kj} G_{kj}(\alpha) + (2\pi/3)\, \rho(r_a = R, r_b = 0), \quad (A17)$$

showing that the second term is proportional to the electron density at the second nucleus b.

References

AVERY, J. & WATSON, K. J. (1977). *Acta Cryst.* A**33**, 679–680

BEVINGTON, P. R. (1969). *Data Reduction and Error Analysis for the Physical Sciences*, pp. 235–240. New York: McGraw Hill.

COPPENS, P. (1977). *Isr. J. Chem.* **16**, 159–162.

DAVIDSON, E. R. & JONES, L. L. (1962). *J. Chem. Phys.* **37**, 2966–2971.

DAWSON, B. (1965). *Aust. J. Chem.* **18**, 595–603.

DAWSON, B. (1967). *Proc. R. Soc. London Ser. A*, **298**, 264–288.

EPSTEIN, J. (1972). BSc Honours Thesis. Univ. of Western Australia.

EPSTEIN, J., BENTLEY, J. & STEWART, R. F. (1977). *J. Chem. Phys.* **66**, 5564–5567.

FINK, M., GREGORY, D. & MOORE, P. G. (1976). *Phys. Rev. Lett.* **37**, 15–18.

FRAGA, S. & RANSIL, B. J. (1961). *J. Chem. Phys.* **35**, 1967–1977.

HANSEN, N. K. & COPPENS, P. (1978). *Acta Cryst.* A**34**, 909–921.

HAREL, M. & HIRSHFELD, F. L. (1975). *Acta Cryst.* B**31**, 162–172.

HEHRE, W. J., STEWART, R. F. & POPLE, J. A. (1969). *J. Chem. Phys.* **51**, 2657–2664.

ISIGURO, E. (1948). *J. Phys. Soc. Jpn*, **3**, 133–135.

KOHL, D. A. & BARTELL, L. S. (1969). *J. Chem. Phys.* **51**, 2891–2904.

KOLOS, W. & ROOTHAAN, C. C. J. (1960). *Rev. Mod. Phys.* **32**, 219–232.

PITZER, R. M., KERN, C. W. & LIPSCOMB, W. N. (1962). *J. Chem. Phys.* **37**, 267–274.

POLITZER, P. (1979). *J. Chem. Phys.* **70**, 1067–1069.

PRICE, P. F. (1976). *The Electron Density in Molecular Crystals.* Thesis, Univ. of Western Australia.

PRICE, P. F., VARGHESE, J. N. & MASLEN, E. N. (1978). *Acta Cryst.* A**34**, 203–216.

REID, R. L. & VAIDA, M. L. (1973). *Phys. Rev. A*, **7**, 1841–1849.

STEVENS, E. D. (1979). *Mol. Phys.* **37**, 27–45.

STEWART, R. F. (1969). *J. Chem. Phys.* **51**, 4569–4577.

STEWART, R. F. (1976). *Acta Cryst.* A**32**, 565–574.

STEWART, R. F. (1977). *Isr. J. Chem.* **16**, 124–131.

STEWART, R. F., BENTLEY, J. & GOODMAN, B. (1975). *J. Chem. Phys.* **63**, 3786–3793.

STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.

VIDAL-VALAT, G., VIDAL, J. P. & KURKI-SUONIO, K. (1978). *Acta Cryst.* A**34**, 594–602.