

Molecular Physics

An International Journal at the Interface Between Chemistry and Physics

ISSN: 0026-8976 (Print) 1362-3028 (Online) Journal homepage: <http://www.tandfonline.com/loi/tmpb20>

An overlap model for estimating the anisotropy of repulsion

Richard J. Wheatley & Sarah L. Price

To cite this article: Richard J. Wheatley & Sarah L. Price (1990) An overlap model for estimating the anisotropy of repulsion, Molecular Physics, 69:3, 507-533, DOI: [10.1080/00268979000100371](https://doi.org/10.1080/00268979000100371)

To link to this article: <https://doi.org/10.1080/00268979000100371>



Published online: 23 Aug 2006.



Submit your article to this journal



Article views: 75



View related articles



Citing articles: 68 View citing articles

An overlap model for estimating the anisotropy of repulsion

by RICHARD J. WHEATLEY and SARAH L. PRICE†

University Chemical Laboratory, Lensfield Road,
Cambridge CB2 1EW, England

(Received 4 July 1989; accepted 21 August 1989)

There is an urgent need for accurate anisotropic site-site intermolecular pair potentials for use in realistic simulations of the condensed phases and spectra of van der Waals molecules. However, all attempts to find an analytical form for the repulsion energy have relied on empirical models with many fitted parameters, whose determination requires large quantities of accurate experimental or *ab initio* data. In this paper, we develop and assess a procedure for estimating the repulsion anisotropy from the monomer wavefunctions which is straightforward, computationally inexpensive and capable of high accuracy. The method is based on the observation that the intermolecular repulsion energy is closely related to the overlap between the unperturbed charge densities of the interacting molecules. The overlap can be treated analytically, leading to an anisotropic site-site functional form. Model pair repulsion potentials with two sites are obtained for $(F_2)_2$ and $(Cl_2)_2$ by ignoring the bonding charge density. For $(N_2)_2$ a novel method of distributing the bonding charge density onto the atoms and the bond centre is used to obtain a three-site model, and this technique is also applied to $(F_2)_2$ and $(Cl_2)_2$. The resulting pair potentials, which have one or two fitted parameters, are found to reproduce an *ab initio* repulsion surface much more accurately than an isotropic atom-atom potential with two parameters. Since the method uses only the monomer charge densities, it can be extended to larger molecules. The resulting repulsion potential is suitable for incorporation into a complete model of the interaction energy.

1. Introduction

A model for the intermolecular forces between polyatomic molecules often constitutes the fundamental scientific input to theoretical studies of the behaviour of the molecules in the solid, liquid or gaseous state. However, the most commonly used kind of model potential, which is the isotropic atom-atom model, has often been found inadequate in realistic simulations; for example, it was unable to predict the anomalous thermal expansivity of solid carbon disulphide [1], the crystal properties of fluorine [2] and the radial distribution function of liquid bromine [3]. The isotropic atom-atom model assumes that the molecules interact as if they are a superposition of spherical charge distributions, which is clearly an approximation, as it ignores the non-spherical distribution of the valence electrons. Thus a more realistic model potential would have an anisotropic site-site form

$$V_{AB}^{\text{model}} = \sum_{a \in A} \sum_{b \in B} V_{ab}^{\text{model}}(R_{ab}, \Omega_{ab}), \quad (1)$$

† Present address: Department of Chemistry, University College, 20 Gordon Street, London WC1H 0AJ, England.

where R_{ab} is the distance from site a on molecule A to site b on molecule B , and Ω_{ab} is the relative orientation of sites a and b , the orientation of a site being defined by the intramolecular bonds, as these determine the orientations of the non-spherical features, such as lone pairs or π -electrons. Experimental evidence that this anisotropy can be important comes from a recent statistical survey [4] of C-X \cdots X-C intermolecular contacts in crystals where, for example, two Br atoms can approach about 0.6 Å more closely when they are 'end-on', with the C-Br \cdots Br angles being 180°, than when they are 'sideways-on', with the C-Br \cdots Br angles being 90°. This effect may be ascribed to the anisotropy arising from the lone pairs. The introduction of anisotropy into the model potential has led to successful simulations of solid and liquid chlorine [5] and the dimer structure of hydrogen fluoride [6]. Including anisotropy in the atom-atom model of the intermolecular pair potential has also been found necessary in attempts to fit *ab initio* potential energy surfaces to an analytical form for (H₂)₂ [7] and (N₂)₂ [8].

The development of accurate model anisotropic atom-atom or site-site potentials involves obtaining both the appropriate functional form for the anisotropy, and values for the large number of parameters involved. This problem has been largely overcome for the long-range forces, where the application of perturbation theory leads to an analytical expression for the electrostatic [9], induction [10] and dispersion [11] energies. However, there is no theoretically justified procedure for obtaining the anisotropy of the repulsion energy accurately, and it is probably no exaggeration to say that this is a major barrier to progress in the field of intermolecular potentials. Determining the form and parameters for the anisotropy of the repulsion by fitting to experimental data is inherently unsatisfactory, as, for example, crystal properties only sample the repulsive wall at a limited number of relative orientations, and other properties usually depend on some form of average over the orientations. Thus, a limited set of experimental data may be reproduced very well by an isotropic or anisotropic atom-atom model potential which extrapolates unrealistically to orientations which are not sampled in the fitting, but which may be important in studying other phases, such as hypothetical crystal structures. In principle, *ab initio* potential energy surfaces do not suffer from this problem, but they are sufficiently difficult to calculate, even for the smallest polyatomics [12], that it is prohibitively expensive to calculate sufficient points on the surface to establish both the form of anisotropy and the parameters unambiguously. An *ab initio* based method of determining the repulsion anisotropy, which is less prohibitive, is the test-particle method of Stone and Tong [13], where fits to the helium-polyatomic surface are used to estimate the polyatomic-polyatomic repulsion potential by the use of combining rules. This is the first non-empirical method in the literature for estimating the anisotropy of the repulsion, and can be used, along with the isotropic atom-atom potential, as a benchmark for other methods.

In this paper, we develop and assess a procedure of estimating the repulsion anisotropy from the monomer wavefunctions which is straightforward, computationally inexpensive and capable of high accuracy. The method is based on the observation [14] that the intermolecular repulsion energy V_{rep} is closely related to the overlap $S_p = \int \rho_A(\mathbf{r})\rho_B(\mathbf{r}) d^3\mathbf{r}$ between the unperturbed charge densities ρ_A , ρ_B of the interacting molecules by a relationship of the form $V_{\text{rep}} \approx K S_p^x$. Previously, the overlap approximation has been applied to the repulsion between pairs of rare gas atoms, and between a rare gas atom and a halide ion [14, 15, 16] with excellent results for $0.96 \leq x \leq 0.99$. It has also been tested at a few points on the N₂-N₂

Need to check this.

surface [16], which gave very good results when x was 0.88. Since the overlap S_ρ can be treated analytically, unlike the repulsion energy V_{rep} , we can develop an analytical expression for the orientation dependence of the repulsion, in terms of the *ab initio* charge densities of the monomers, by assuming this relationship. The resulting expression is an anisotropic site-site model for the repulsion potential. It has only one parameter K , if we assume a value for x . K needs to be determined by reference to experimental data or an *ab initio* potential energy surface.

The first stage in the development of the overlap model is to derive analytical expressions for the orientation dependence of the overlaps involving two orbitals on the same atom in molecule A , and two orbitals on the same atom in molecule B , for linear molecules described using *sp* basis sets. This illustrates the general approach, and the dependence of the form of the anisotropy on the valence electron distribution. This atomic overlap model $S_{\rho, \text{atomic}}$ will be applied to give an anisotropic atom-atom repulsion model for F_2 and Cl_2 , which are molecules with very little bonding electron density. However, the full overlap S_ρ also includes 3 and 4 centre terms, arising from the terms in the charge density of the molecules involving orbitals on different centres. These are treated in § 3, by a novel method of representing the charge density in terms of gaussian type functions at atom and bond centre sites. This enables the expressions derived in § 2 to be applied to give a 3 site anisotropic model $S_{\rho, \text{dist}}$ for the repulsion interaction of diatomic molecules where there is significant bonding density, such as N_2 . These models are tested against *ab initio* repulsion potential energy surfaces for $(\text{F}_2)_2$, $(\text{Cl}_2)_2$ and $(\text{N}_2)_2$. We shall see that the overlap model can predict these surfaces with approximately twice the accuracy of an isotropic atom-atom model, with fewer fitted parameters, and indeed compare favourably with the best available unconstrained fits to these potential energy surfaces.

2. The atomic overlap model

The *ab initio* charge density ρ_A of a molecule A is given by the equation

$$\rho_A(\mathbf{r}) = \sum_{i,j} \phi_i(\mathbf{r})\phi_j(\mathbf{r})M_{ij}. \quad (2)$$

M_{ij} is a density matrix element, and ϕ_i and ϕ_j are atomic orbitals of A . If standard *ab initio* techniques [17] are used, then the atomic orbitals will be linear combinations of gaussian type functions. A cartesian form for the gaussians is usually adopted

$$\phi_i(\mathbf{r}) = \sum_u a_{iu} \eta_u(\mathbf{r}) = \sum_u a_{iu} (x - x_i)^{\lambda_u} (y - y_i)^{\mu_u} (z - z_i)^{\nu_u} \exp [-\alpha_u (\mathbf{r} - \mathbf{r}_i)^2], \quad (3)$$

where η_u is a gaussian type orbital centred at $\mathbf{r}_i = (x_i, y_i, z_i)$, and a_{iu} is the product of a normalization constant and a contraction coefficient.

The rank, l_u , of η_u is defined as $\lambda_u + \mu_u + \nu_u$. The product of a pair of gaussians η_u and η_v , of ranks l_u and l_v , is a sum of gaussians, of ranks 0, 1, 2, ..., $l_u + l_v$, centred at a point on the line joining the centres of η_u and η_v , whose position is determined by the exponents of the gaussians (see Appendix). The charge density of an *ab initio* wavefunction can be expressed exactly as a sum of gaussians centred at a large number of *overlap sites* \mathbf{r}_n :

$$\rho_A(\mathbf{r}) = \sum_{n \in A} \sum_w k_w (x - x_n)^{\lambda_w} (y - y_n)^{\mu_w} (z - z_n)^{\nu_w} \exp [-\alpha_w (\mathbf{r} - \mathbf{r}_n)^2], \quad (4)$$

where k_w contains many constants including the density matrix elements (see

Appendix for detailed form), and $\alpha_w = \alpha_u + \alpha_v$. If the gaussians comprising the wavefunction are all centred on the atoms, then the overlap sites will all be on lines joining the atoms. This complete expression will be considered in the next section, but first we shall consider only those terms where ϕ_i and ϕ_j are on the same atom, so all the overlap sites \mathbf{r}_n are on the atoms.

The overlap $S_{\rho, \text{atomic}}$ between two of these atomic charge densities is given by a sum of overlaps between gaussians, and hence an analytical expression for $S_{\rho, \text{atomic}}$ can be obtained. The following example illustrates the method of calculation. Consider two molecules A and B . We shall perform the calculation in the coordinate system of A , and define the relative orientation of the sites in terms of the unit vectors $\hat{\mathbf{z}}_A = (0, 0, 1)$, and $\hat{\mathbf{z}}_B = (\hat{\mathbf{x}}_A \cdot \hat{\mathbf{z}}_B, \hat{\mathbf{y}}_A \cdot \hat{\mathbf{z}}_B, \hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)$, where $\hat{\mathbf{z}}_A$ and $\hat{\mathbf{z}}_B$ are unit vectors along the molecular bond which point towards the other atom in the same molecule. We shall evaluate the overlap between a z -type gaussian $\eta_z^a(\mathbf{r})$ on site a , and another z -type gaussian $\eta_z^b(\mathbf{r})$ on site b . Such gaussians would arise, for example, from the product of an s and a p_z gaussian basis function on the atom. It is necessary to re-express η_z^b in the coordinate system of A , which is achieved by using the above expression for $\hat{\mathbf{z}}_B$. Site a is assumed to be at $(0, 0, 0)$, and site b will then be at $(\hat{\mathbf{x}}_A \cdot \mathbf{R}, \hat{\mathbf{y}}_A \cdot \mathbf{R}, \hat{\mathbf{z}}_A \cdot \mathbf{R})$, where \mathbf{R} is the vector from a to b . The gaussians have the form

$$\eta_z^a(\mathbf{r}) = k_a z \exp [-\alpha_a r^2] \quad (5a)$$

and

$$\begin{aligned} \eta_z^b(\mathbf{r}) = k_b [& (x - \hat{\mathbf{x}}_A \cdot \mathbf{R}) \hat{\mathbf{x}}_A \cdot \hat{\mathbf{z}}_B + (y - \hat{\mathbf{y}}_A \cdot \mathbf{R}) \hat{\mathbf{y}}_A \cdot \hat{\mathbf{z}}_B + (z - \hat{\mathbf{z}}_A \cdot \mathbf{R}) \hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B] \\ & \times \exp \{ -\alpha_b [(x - \hat{\mathbf{x}}_A \cdot \mathbf{R})^2 + (y - \hat{\mathbf{y}}_A \cdot \mathbf{R})^2 + (z - \hat{\mathbf{z}}_A \cdot \mathbf{R})^2] \}. \end{aligned} \quad (5b)$$

We define s_ρ to be a single overlap between two gaussians, so, from our equation for the overlap,

$$\begin{aligned} s_\rho(\mathbf{R}, \Omega; \eta_z^a, \eta_z^b) &= k_a k_b \int z [(x - \hat{\mathbf{x}}_A \cdot \mathbf{R}) \hat{\mathbf{x}}_A \cdot \hat{\mathbf{z}}_B + (y - \hat{\mathbf{y}}_A \cdot \mathbf{R}) \hat{\mathbf{y}}_A \cdot \hat{\mathbf{z}}_B + (z - \hat{\mathbf{z}}_A \cdot \mathbf{R}) \hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B] \\ &\quad \times \exp [-\alpha_a x^2 - \alpha_b x^2 + 2\alpha_b x \hat{\mathbf{x}}_A \cdot \mathbf{R} - \alpha_b (\hat{\mathbf{x}}_A \cdot \mathbf{R})^2] \\ &\quad \times \exp [-\alpha_a y^2 - \alpha_b y^2 + 2\alpha_b y \hat{\mathbf{y}}_A \cdot \mathbf{R} - \alpha_b (\hat{\mathbf{y}}_A \cdot \mathbf{R})^2] \\ &\quad \times \exp [-\alpha_a z^2 - \alpha_b z^2 + 2\alpha_b z \hat{\mathbf{z}}_A \cdot \mathbf{R} - \alpha_b (\hat{\mathbf{z}}_A \cdot \mathbf{R})^2] dx dy dz. \end{aligned} \quad (6)$$

To integrate this, we complete the square in the exponent and introduce a change of variables

$$\zeta_x = x - \frac{\alpha_b}{\alpha_a + \alpha_b} \hat{\mathbf{x}}_A \cdot \mathbf{R}, \quad \zeta_y = y - \frac{\alpha_b}{\alpha_a + \alpha_b} \hat{\mathbf{y}}_A \cdot \mathbf{R}, \quad \zeta_z = z - \frac{\alpha_b}{\alpha_a + \alpha_b} \hat{\mathbf{z}}_A \cdot \mathbf{R}. \quad (7)$$

This gives

$$\begin{aligned} s_\rho(\mathbf{R}, \Omega; \eta_z^a, \eta_z^b) &= k_a k_b \exp \left[-\frac{\alpha_a \alpha_b}{\alpha_a + \alpha_b} \mathbf{R}^2 \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\zeta_z + \frac{\alpha_b}{\alpha_a + \alpha_b} \hat{\mathbf{z}}_A \cdot \mathbf{R} \right) \\ &\quad \times \left[\left(\zeta_x - \frac{\alpha_a}{\alpha_a + \alpha_b} \hat{\mathbf{x}}_A \cdot \mathbf{R} \right) \hat{\mathbf{x}}_A \cdot \hat{\mathbf{z}}_B + \left(\zeta_y - \frac{\alpha_a}{\alpha_a + \alpha_b} \hat{\mathbf{y}}_A \cdot \mathbf{R} \right) \hat{\mathbf{y}}_A \cdot \hat{\mathbf{z}}_B \right. \\ &\quad \left. + \left(\zeta_z - \frac{\alpha_a}{\alpha_a + \alpha_b} \hat{\mathbf{z}}_A \cdot \mathbf{R} \right) \hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B \right] \\ &\quad \times \exp [-(\alpha_a + \alpha_b)(\zeta_x^2 + \zeta_y^2 + \zeta_z^2)] d\zeta_x d\zeta_y d\zeta_z, \end{aligned} \quad (8)$$

where we have used $R^2 = (\hat{\mathbf{x}}_A \cdot \mathbf{R})^2 + (\hat{\mathbf{y}}_A \cdot \mathbf{R})^2 + (\hat{\mathbf{z}}_A \cdot \mathbf{R})^2$. We can now integrate equation (8) using the standard integrals

$$\begin{aligned} \int_{-\infty}^{\infty} \exp(-\alpha \zeta_x^2) d\zeta_x &= \sqrt{\frac{\pi}{\alpha}}, \quad \int_{-\infty}^{\infty} \zeta_x \exp(-\alpha \zeta_x^2) d\zeta_x = 0, \\ \int_{-\infty}^{\infty} \zeta_x^2 \exp(-\alpha \zeta_x^2) d\zeta_x &= \frac{1}{2\alpha} \sqrt{\frac{\pi}{\alpha}}. \end{aligned} \quad (9)$$

The result is

$$s_\rho(R, \Omega; \eta_z^a, \eta_z^b) = k_a k_b \exp \left[-\frac{\alpha_a \alpha_b}{\alpha_a + \alpha_b} R^2 \right] \times \left\{ \frac{1}{2} \sqrt{\left[\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right]} \hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B \right. \\ \left. - \alpha_a \alpha_b R^2 \sqrt{\left[\frac{\pi^3}{(\alpha_a + \alpha_b)^7} \right]} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}) \right\}, \quad (10)$$

where we have inserted $\hat{\mathbf{z}}_B \cdot \mathbf{R}$ in place of its expansion $[(\hat{\mathbf{x}}_A \cdot \hat{\mathbf{z}}_B)(\hat{\mathbf{x}}_A \cdot \mathbf{R}) + (\hat{\mathbf{y}}_A \cdot \hat{\mathbf{z}}_B)(\hat{\mathbf{y}}_A \cdot \mathbf{R}) + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)(\hat{\mathbf{z}}_A \cdot \mathbf{R})]$ in the coordinate system of *A* and replaced \mathbf{R} by $R\hat{\mathbf{R}}$, so that the orientation dependence is expressed in terms of direction cosines.

All the s_ρ are evaluated in this way, and then added together to give $S_{\rho, \text{atomic}}$. This task is considerably easier if the interacting molecules are linear, in which case most of the k_w are zero, because the density matrix elements between *s* and p_x , *s* and p_y , p_x and p_y , p_x and p_z , and p_y and p_z orbitals are all zero by symmetry. This means that λ_w and μ_w in equation (4) must both be even for k_w to be non-zero. If an *sp* basis set is used, then the maximum rank of the gaussians comprising the atomic charge density is two. The explicit form of all possible overlaps between these gaussian types is given in the Appendix. When all the contributing terms are added together, an expression of the following form is obtained for the overlap in terms of its orientation dependence:

$$\begin{aligned} S_{\rho, \text{atomic}}(R, \Omega) &= \sum_{a \in A} \sum_{b \in B} c_{000}^{ab}(R_{ab}) + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})c_{100}^{ab}(R_{ab}) + (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})c_{010}^{ab}(R_{ab}) \\ &\quad + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)c_{001}^{ab}(R_{ab}) + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})^2 c_{200}^{ab}(R_{ab}) + (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})^2 c_{020}^{ab}(R_{ab}) \\ &\quad + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)^2 c_{002}^{ab}(R_{ab}) + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})c_{110}^{ab}(R_{ab}) \\ &\quad + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)c_{101}^{ab}(R_{ab}) + (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)c_{011}^{ab}(R_{ab}) \\ &\quad + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})^2 (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})c_{210}^{ab}(R_{ab}) + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})^2 c_{120}^{ab}(R_{ab}) \\ &\quad + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})^2 (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})^2 c_{220}^{ab}(R_{ab}) \\ &\quad + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)c_{111}^{ab}(R_{ab}), \end{aligned} \quad (11)$$

$$= \sum_{a \in A} \sum_{b \in B} \sum_{ijk} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab})^i (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab})^j (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)^k c_{ijk}^{ab}(R_{ab}) \quad (11a)$$

where

$$i + k \leq 2 \text{ and } j + k \leq 2.$$

The anisotropy coefficients c_{ijk}^{ab} contain all the radial dependence of the overlap, and themselves depend on the parameters of the monomer wavefunction.

We have calculated all the anisotropy coefficients for $(F_2)_2$ and $(Cl_2)_2$ for several fixed atom–atom separations R_{ab} , using parameters from *ab initio* wavefunctions. The CADPAC [18] suite of programs was used to obtain the monomer wavefunctions, using standard basis sets taken from the CADPAC library. For fluorine, a double-zeta basis was used, with a bondlength of $2.678 a_0$, and for chlorine, a 6-31 G* basis was used, with a bondlength of $3.757 a_0$. For chlorine, the overlap model had to be extended to include *d* functions, which was a straightforward but lengthy process, and has been included in the wavefunction analysis computer program.

Calculation of the anisotropy coefficients at all separations would give an exact value for the atomic overlap at any orientation, but in order to obtain a functional form which is suitable for use in a model potential, it is necessary to remove the separate R_{ab} dependence of the c_{ijk}^{ab} . This involves introducing two further approximations. Firstly, we assume that the ratios $c_{ijk}^{ab}/c_{000}^{ab}$ are independent of R_{ab} ; this assumption is tested in figure 1 for fluorine and chlorine. The variation of the smaller ratios is negligible, and the larger parameters are very nearly constant; they vary in magnitude by up to 3 per cent of c_{000} from their values at the van der Waals diameter R_{vdw} of the atoms, which is $5.1 a_0$ for fluorine and $6.8 a_0$ for chlorine [19]. We shall use the values of the parameters at the van der Waals diameter. This choice is to some extent arbitrary, but it is unlikely to introduce a significant error, since the ratios vary little with R_{ab} . Secondly, since the repulsion energy depends approximately exponentially on separation [20], and since the overlap between the charge density arising from Slater (or hydrogen-like) orbitals is also approximately exponential [16], we expect that the c_{000}^{ab} should also decay exponentially with R_{ab} for a reasonable monomer wavefunction. This is tested in figure 2, which shows clearly that the exponential assumption is extremely good for regions within $\pm 0.5 a_0$ ($\pm 0.25 \text{ \AA}$) of the van der Waals diameter. These approximations lead to the expression for the atomic overlap which will be used in our model potentials:

$$S_{\rho, \text{atomic}} = \sum_{a \in A} \sum_{b \in B} C_{000}^{ab} \exp(-\alpha^{ab} R_{ab}) [1 + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}_{ab}) C_{100}^{ab} + (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}_{ab}) C_{010}^{ab} + \dots], \quad (12)$$

where the C_{ijk}^{ab} , defined by $C_{ijk}^{ab} = c_{ijk}^{ab}/c_{000}^{ab}$ for $ijk \neq 000$, are now assumed independent of R_{ab} , and the α^{ab} , given by $\alpha^{ab} R_{ab} = \ln C_{000}^{ab} - \ln c_{000}^{ab}$, describe the exponential radial dependence of the overlap. We are not interested in the absolute value of the overlap, only in quantities which are proportional to it, so we define $C_{000}^{\text{at-at}}$, the atom–atom C_{000}^{ab} , to be 1.0. The parameters C and exponents α for the $(F_2)_2$ and $(Cl_2)_2$ systems, calculated at R_{vdw} , are summarized in table 1. The exponents α were obtained from the gradients of the straight lines in figure 2.

We know from elementary molecular orbital theory that the fluorine and chlorine atoms are quite anisotropic within the molecules, resembling oblate ellipsoids due to the lone pairs perpendicular to the molecular axes. This results in large $\eta_{x^2+y^2}$ terms in the charge densities, which should give large negative $(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2$ and $(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2$ terms and a large positive $(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2$ term in the overlap (see Appendix). Table 1 confirms this: C_{200} , C_{020} and C_{220} are the largest parameters, and they have the correct signs. The repulsion around the van der Waals diameter is almost entirely due to the valence electrons, since the factor of $\exp[-\alpha_a \alpha_b R_{ab}^2 / (\alpha_a + \alpha_b)]$ which appears in equation (A 5) is very small for overlaps involving core orbitals, as they have large exponents. Chlorine is consistently more anisotropic

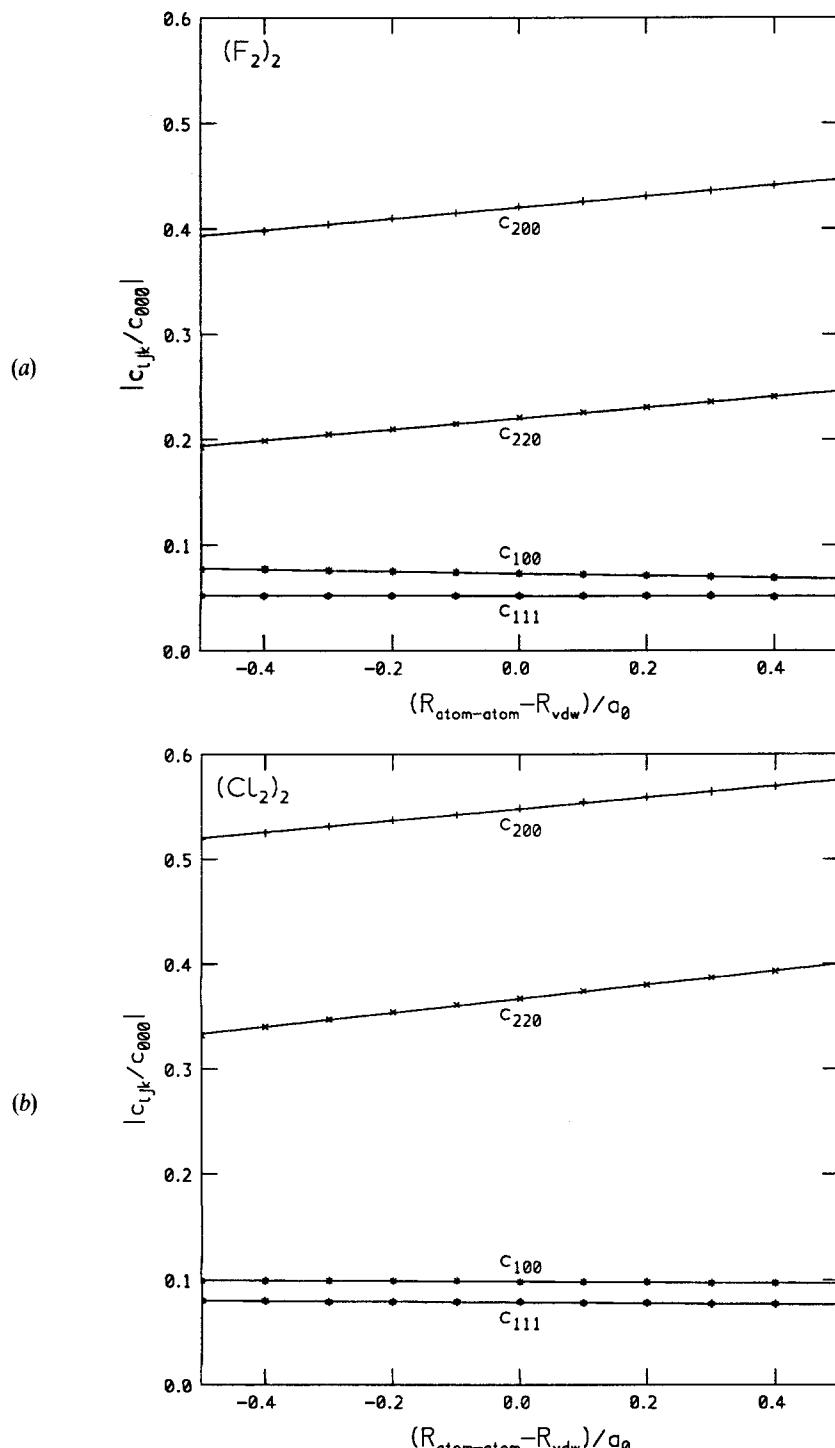


Figure 1. R -dependence of the anisotropy parameters. The four largest anisotropy parameters for the atomic overlap of $(\text{F}_2)_2$ (a) and $(\text{Cl}_2)_2$, (b) are plotted as a function of separation $R_{\text{atom-atom}}$. We used $R_{\text{vdw}} = 5 \cdot a_0 = 2.7 \text{ \AA}$ for fluorine, and $6.8 \cdot a_0 = 3.6 \text{ \AA}$ for chlorine as estimates of the atomic van der Waals diameters.

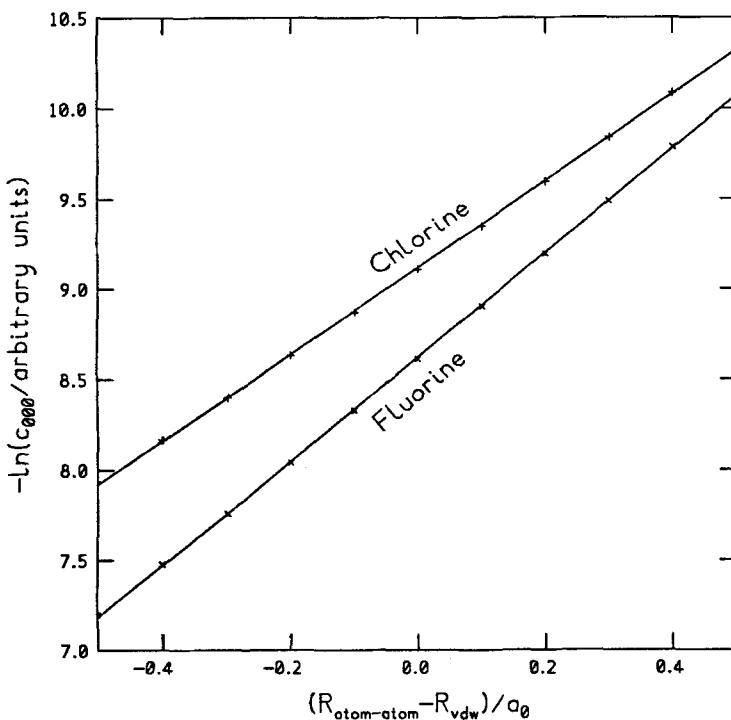


Figure 2. Investigation of the exponential assumption. The c_{000} parameter for the atomic overlap of $(F_2)_2$ and $(Cl_2)_2$ is shown to be exponential to a very high degree of accuracy. The best-fit lines have correlation coefficients of 0.99994 (for fluorine) and 0.99993 (for chlorine).

than fluorine, as expected on the basis of the crystal structures [21], but the anisotropy parameters have similar relative values for the two systems. This indicates that the atomic overlap model could be transferable between chemically similar species.

The obvious extension to the atomic overlap model is to consider the terms arising from the bonding electron density, which appeared in equation (4) but were then ignored. We have found that a method which distributes the charge density to

Table 1. Atomic overlap parameters for F_2-F_2 and Cl_2-Cl_2 .

	F_2-F_2	Cl_2-Cl_2
α/a_0^{-1}	2.89	2.40
C_{000}	1.000†	1.000†
$C_{100} = -C_{010}$	-0.073	-0.098
$C_{200} = C_{020}$	-0.421	-0.548
$C_{210} = -C_{120}$	-0.038	-0.062
C_{220}	0.221	0.367
C_{111}	-0.052	-0.079

† By definition.

Parameters less than 0.02 have been ignored. This includes all extra parameters arising from the d functions in the chlorine basis set, that is, ones which are not included in equation (11).

the atoms and bond centres works well, and we discuss this in the next section. In § 4, we test both the atomic and the distributed overlap models.

3. The distributed overlap model

3.1. Distributing the charge density

In the last section, we showed how to obtain an expression for the overlap between a pair of gaussian type functions at an arbitrary separation and orientation. Since we can express the charge density of an *ab initio* wavefunction exactly in terms of gaussians, we can obtain an exact analytical expression for the overlap between the charge densities of two wavefunctions. However, this expression would contain a separate term for the interaction of each overlap site on one molecule with each overlap site on the other. This would make the total number of terms unacceptably large.

Our solution is to *distribute* the charge density of the wavefunctions; in other words, to express it as a sum of gaussians located at a relatively small number of *expansion centres*, such as the atoms and the bond centres. It is necessary to move each gaussian in turn from an overlap site to an expansion centre. If the overlap site is itself an expansion centre, then the gaussian is unchanged. If not, we move the gaussian to the nearest expansion centre.

The formula for moving a gaussian is derived as follows. Consider a typical gaussian $k_w(x - x_n)^{\lambda_w}(y - y_n)^{\mu_w}(z - z_n)^{\nu_w} \exp[-\alpha(r - r_n)^2]$. We wish to move it from the overlap centre r_n to an expansion centre r_p . Let $r_v = r_p - r_n$. Then

$$\begin{aligned} \exp[-\alpha(r - r_n)^2] &= \exp[-\alpha(r - r_p + r_v)^2] \\ &= \exp[-\alpha r_v^2] \exp[-\alpha(r - r_p)^2] \exp[-2\alpha(r - r_p) \cdot r_v] \\ &= \exp[-\alpha r_v^2] \exp[-\alpha(r - r_p)^2] \sum_{L=0}^{\infty} \frac{(-2\alpha(r - r_p) \cdot r_v)^L}{L!}. \quad (13) \end{aligned}$$

Of the three terms in this product, the first is a constant, and the second is a spherical gaussian centred at r_p . Since

$$(r - r_p) \cdot r_v = x_v(x - x_p) + y_v(y - y_p) + z_v(z - z_p),$$

the third term is a sum of functions, each of the form $\kappa(x - x_p)^\lambda(y - y_p)^\mu(z - z_p)^\nu$. The pre-exponential part of the original gaussian may be re-expressed more straightforwardly, since, for example, $(x - x_n)^{\lambda_w} = (x - x_p + x_v)^{\lambda_w}$, and a binomial expansion gives this term as a sum of powers of $(x - x_p)$, as required.

Thus, the original gaussian has been re-expressed in terms of an infinite sum of gaussians centred at the expansion centre. In the event that the overlap centre is equidistant from two or more expansion centres, we divide the gaussian into the appropriate number of identical parts, and move each part to a different one of these expansion centres.

The bonding charge density has now been distributed to a small number of sites, while the atomic charge density has remained constant. The distributing procedure leads to an infinite series at each site, which contains the expansion of an exponential involving r_v . If atoms and bond centres are used as expansion centres, then r_v is typically only half as large as in the case when only the atoms are used as the expansion centres, and the convergence is expected to be much better. The rate of

convergence is very important, because it is necessary to truncate the infinite series by ignoring all terms with rank greater than some maximum rank l_{\max} . The accuracy of the expansion has been tested for five diatomics by comparing the distributed and true *ab initio* charge densities at evenly spaced points on a rectangular grid. Calculations have been performed using a 3-site distributed model with $l_{\max} = 4$, a 3-site distributed model with $l_{\max} = 2$, a 2-site distributed model with $l_{\max} = 4$ and the atomic charge density, as defined in § 2, which completely ignores the bonding density. Table 2 summarizes the results. It is clearly necessary, in general, to have expansion centres at the bond centres as well as at the nuclei, and to have $l_{\max} = 4$. However, this gives an excellent representation, with an rms error of 1 per cent or better. It is interesting that the atomic model is superior to a distributed 2-site model in most cases; clearly ignoring the bonding density can be a better approximation than distributing it to the atoms using only 5 terms. For fluorine and chlorine, the atomic expansion represents the charge density well and so could give sufficient accuracy to be used in a model potential. For the other systems, where the bonding electron density is much higher compared to the atomic electron density, the atomic expansion would not be suitable.

It may also be useful to see where the distributed model is most accurate, so we have compared the distributed charge density of nitrogen (calculated using the 3-site, $l_{\max} = 4$ expansion) with the true *ab initio* charge density at a large number of evenly spaced points. The results are plotted in figure 3. We see that the model is very accurate at long range, and has small areas of 2 per cent inaccuracy at the bond centre (which is unimportant), and at the ends of the molecule (which may be significant). The charge density is quite exponential out to the van der Waals radius, since the contours in figure 3(a) are almost equally spaced, confirming our belief that the radial part of the overlap can also be represented by an exponential decay.

Table 2. Comparison of distributed and *ab initio* charge densities.

Molecule	Orientation	x_{\max} $/a_0$	z_{\max} $/a_0$	Number of points	R.M.S. percentage difference		
					Atomic	Distributed	
Fluorine	F at $(0, 0, \pm 1.339)$	6.0	6.0	961	4.61	8.93	0.95
Chlorine	Cl at $(0, 0, \pm 1.8785)$	7.0	7.0	676	6.41	7.31	3.59
Nitrogen	N at $(0, 0, \pm 1.0335)$	6.0	6.0	676	15.62	18.29	2.74
HF	H at $(0, 0, 6.739)$ F at $(0, 0, 5.0)$	5.0	10.0	961	18.73	10.76	0.58
CO	C at $(0, 0, 8.132)$ O at $(0, 0, 6.0)$	6.0	12.0	676	17.14	595.21†	7.68
							1.10

† This result reflects unusually large errors in the charge density in the region $5-6 a_0$ beyond the oxygen end of the molecule.

The percentage difference between the distributed and *ab initio* charge density was measured at each point on a rectangular grid in the xz plane, with opposite corners at $(0, 0, 0)$ and $(x_{\max}, 0, z_{\max})$. There were either 26 or 31 points equally spaced along each side of the grid, giving rise to 676 or 961 points altogether.

The basis sets used were double-zeta for F₂, 6-31 G* for Cl₂, 6-311 G* for CO and N₂, and 8s6p for HF.

The atomic model has $l_{\max} = 2$ for the charge densities which were calculated using *sp* basis sets (fluorine and hydrogen fluoride), and $l_{\max} = 4$ for the others. These values give an exact expansion of the atomic charge density.

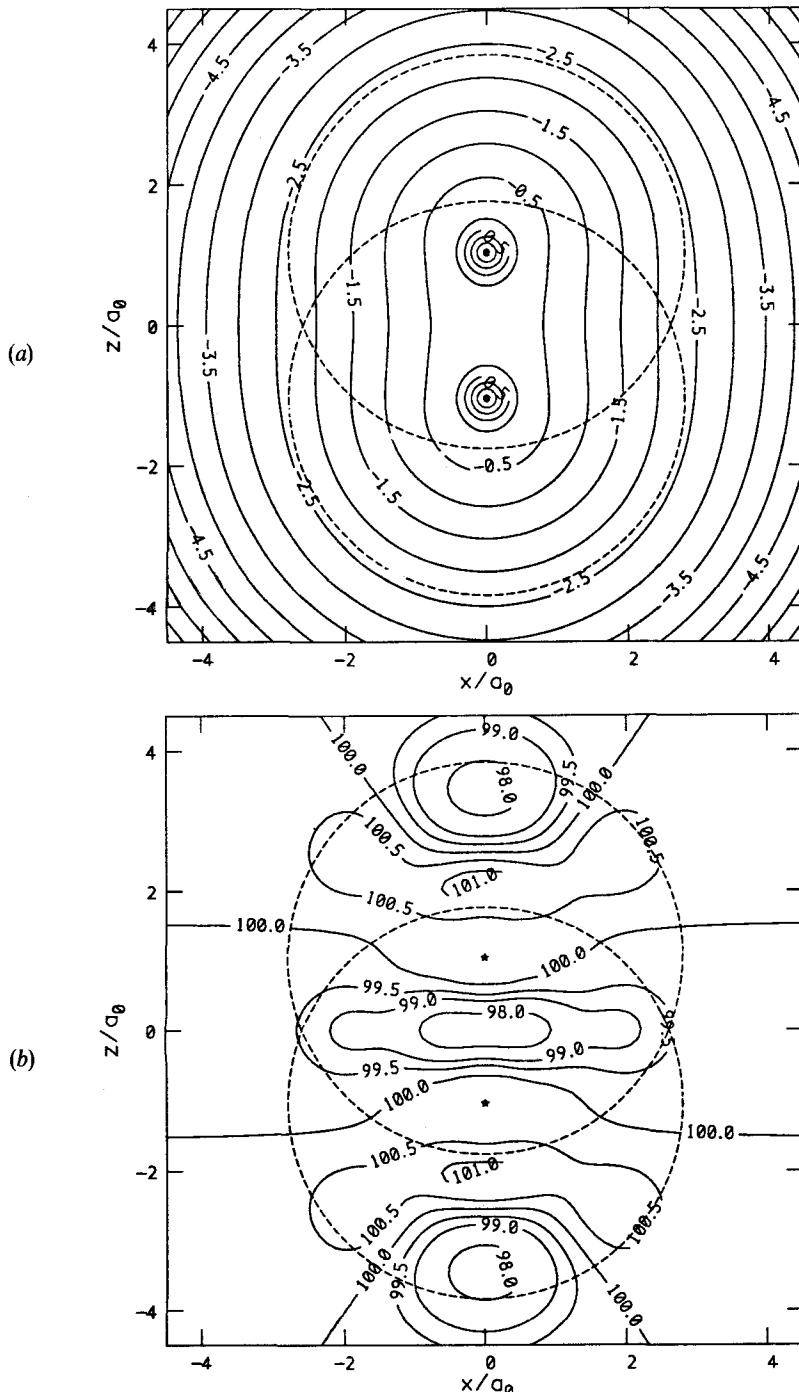


Figure 3. Distributed charge density of nitrogen. In (a), the logarithm of the distributed charge density (in arbitrary units) is shown with contours drawn at equal intervals. The true *ab initio* charge density gives a plot which is exactly superimposable on this scale. In (b), the distributed charge density of nitrogen is plotted as a percentage of the *ab initio* charge density. Contours are drawn at 98.0, 99.0, 99.5, 100.0, 100.5, and 101.0 per cent. Nowhere did the percentage reach as low as 97, or as high as 102 per cent. The atomic van der Waals radii are represented by dashed circles centred on the nuclei, which are shown by stars.

The benefits of distributing the charge density in this way are considerable. It is a highly accurate method, with an error of much less than 1 per cent in most cases. The expression which is obtained for the distributed charge density is in exactly the same form as equation 4, with a dramatic reduction in the number of sites, so the overlap model may be applied to the distributed charge density in exactly the same way as in § 2, except that atom–bond and bond–bond terms will also appear. The implementation of the overlap model is discussed next, and the results it yields are presented in § 4.

3.2. Implementation of the overlap model

In § 2, we showed how to obtain an exact expression for the overlap between two atomic charge densities, and how to approximate this in order to obtain a more convenient functional form. We shall now discuss, in more detail, the procedure which we used to obtain the overlap between the distributed charge densities of fluorine, chlorine and nitrogen.

The first step was to obtain an *ab initio* wavefunction for the monomer. This was the most computationally expensive step, since the time taken for an SCF calculation depends approximately on the fourth power of the number of basis functions. The basis sets and geometries are given in § 2 and table 2. A three-site distributed charge density expansion (§ 3.1) was then performed on the monomers, with $l_{\max} = 4$. This re-expresses the charge density as a sum of gaussians, with ranks from 0 to 4, centred on the atoms and the bond centre. We have already seen that this approximation is within 1 per cent of the true *ab initio* charge density. The time taken to distribute the charge density depends only on the square of the number of basis functions.

The site–site overlaps were then calculated. There are three distinct kinds of term for each interaction: four atom–atom terms, four atom–bond terms and one bond–bond term. Thus, three calculations had to be performed on each molecule. Each calculation was performed in exactly the same way as the calculation of the atom–atom overlap parameters, and gave one α^{ab} value and one set of c_{ijk}^{ab} values. The separations chosen were $5.1 a_0$ for fluorine, $5.7 a_0$ for nitrogen and $6.8 a_0$ for chlorine, as in § 2. The parameter ratios are fairly insensitive of the choice of R_{ab} , so this choice is not vital. The C_{ijk}^{ab} were then obtained. The form of equation (12) means that we can arbitrarily fix one of the C_{000}^{ab} , and we have chosen $C_{000}^{\text{at-at}} = 1.0$, as for the atomic model. The atom–bond centre $C_{000}^{\text{at-bc}}$ and bond centre–bond centre $C_{000}^{\text{bc-bc}}$ were obtained using the following equation, which is derived by comparing terms in equations (11) and (12):

$$\frac{C_{000}^{ab} \exp(-\alpha^{ab} R_{\text{vdw}})}{C_{000}^{\text{at-at}} \exp(-\alpha^{\text{at-at}} R_{\text{vdw}})} = \frac{c_{000}^{ab}(R_{\text{vdw}})}{c_{000}^{\text{at-at}}(R_{\text{vdw}})}$$

so

$$C_{000}^{ab} = \frac{c_{000}^{ab}(R_{\text{vdw}})}{c_{000}^{\text{at-at}}(R_{\text{vdw}})} \exp [(\alpha^{ab} - \alpha^{\text{at-at}})R_{\text{vdw}}]. \quad (14)$$

The anisotropy parameters C_{ijk}^{ab} were obtained using

$$C_{ijk}^{ab} = \frac{c_{ijk}^{ab}(R_{\text{vdw}})}{c_{000}^{ab}(R_{\text{vdw}})}, \quad (15)$$

which is obtained in the same way as equation (14). All the C and α parameters are listed in table 3. For fluorine and chlorine, the largest atom–atom parameters are much more important than the atom–bond parameters, since $C_{000}^{\text{at}-\text{bc}}$ is so small. They are close in value to the atom–atom parameters obtained using the atomic overlap model (table 1). For nitrogen, the atom–bond parameters are quite important, as we would expect, and their anisotropy is very significant. We can still understand the atom–atom anisotropy parameters qualitatively: the large density matrix element between the valence s and p_z orbitals of nitrogen leads to a large η_z gaussian in the charge distribution, and the overlaps between two of these, and between η_z and η_s gaussians, give rise to the large $C_{100}^{\text{at}-\text{at}}$, $C_{010}^{\text{at}-\text{at}}$ and $C_{110}^{\text{at}-\text{at}}$ terms (see Appendix). These shift the repulsive wall outwards at the ends of the molecule, in contrast to the anisotropy of the halogens, where the repulsive wall is shifted inwards. A potential for nitrogen fitted to an *ab initio* surface [8] also predicts this effect.

Table 3. Overlap parameters for use in 3-site model potentials.

	F_2-F_2	Cl_2-Cl_2	N_2-N_2
Overlap parameters between two atoms			
$\alpha^{\text{at}-\text{at}}/a_0^{-1}\dagger$	2.89	2.40	2.47
$\alpha_{OM3}^{\text{at}-\text{at}}/a_0^{-1}\ddagger$	2.46	1.92	2.10
$C_{000}^{\text{at}-\text{at}}\dagger$	1.000	1.000	1.000
$C_{100}^{\text{at}-\text{at}} = -C_{010}^{\text{at}-\text{at}}$	-0.063	-0.098	-0.466
$C_{200}^{\text{at}-\text{at}} = C_{020}^{\text{at}-\text{at}}$	-0.404	-0.546	
$C_{110}^{\text{at}-\text{at}}$			-0.219
$C_{210}^{\text{at}-\text{at}} = -C_{120}^{\text{at}-\text{at}}$	-0.032	-0.062	
$C_{220}^{\text{at}-\text{at}}$	0.199	0.363	
$C_{111}^{\text{at}-\text{at}}$	-0.048	-0.078	
Overlap parameters between an atom and a bond centre			
$\alpha^{\text{at}-\text{bc}}/a_0^{-1}\dagger$	2.61	2.18	2.25
$\alpha_{OM3}^{\text{at}-\text{bc}}/a_0^{-1}\ddagger$	2.22	1.74	1.91
$C_{000}^{\text{at}-\text{bc}}\dagger$	-0.0179	-0.0175	0.121
$C_{000}^{\text{at}-\text{bc}}\ddagger$	-0.0217	-0.0235	0.149
$C_{100}^{\text{at}-\text{bc}}$			-0.389
$C_{200}^{\text{at}-\text{bc}}$	-0.344	-0.488	
$C_{020}^{\text{at}-\text{bc}}$	1.765	0.479	-1.793
$C_{120}^{\text{at}-\text{bc}}$			-1.306
$C_{220}^{\text{at}-\text{bc}}$	-0.829	-0.342	
Overlap parameters between two bond centres			
$\alpha^{\text{bc}-\text{bc}}/a_0^{-1}\dagger$		1.86	
$\alpha_{OM3}^{\text{bc}-\text{bc}}/a_0^{-1}\ddagger$		1.58	
$C_{000}^{\text{bc}-\text{bc}}\dagger$		0.0039	
$C_{000}^{\text{bc}-\text{bc}}\ddagger$		0.0064	
$C_{200}^{\text{bc}-\text{bc}} = C_{020}^{\text{bc}-\text{bc}}$		-1.785	
$C_{220}^{\text{bc}-\text{bc}}$		4.494	
$C_{111}^{\text{bc}-\text{bc}}$		-1.306	

† These parameters are the ones defined in § 3.2, and were used for OM1 and OM2.

‡ These parameters have been modified as explained in § 4, and were used for OM3. All other parameters were used for OM1, OM2 and OM3.

Parameters are included on the basis of their importance relative to $C_{000}^{\text{at}-\text{at}}$ at the van der Waals diameter R_{vdw} . This is given by $C_{ijk}^{ab} C_{000}^{ab} \exp[-R_{\text{vdw}}(\alpha^{ab} - \alpha^{\text{at}-\text{at}})]$ for the parameter C_{ijk}^{ab} . The cutoff values are 0.02 for fluorine and chlorine, and 0.1 for nitrogen. These were chosen to give a reasonably small number of parameters, and such that adding further parameters made a negligible difference to the errors in fitting.

In this section and §2 we have obtained the convenient functional forms $S_{\rho, \text{atomic}}$ and $S_{\rho, \text{dist}}$ for the overlap, which may be calculated much more quickly than the exact overlap. It is also necessary to check that these distributed overlap values are close to the true overlap values. This has been done for fluorine in figure 4. The true overlap values were calculated by integrating over all sets of four gaussian basis functions (two on A and two on B). The time taken for each of these calculations was similar to that for an SCF monomer calculation, so it was only practicable to do it for fluorine. The distributed overlap values were given by equation (12), using only the parameters in table 3. The agreement is very good, so our model potential closely resembles the true overlap. In the next section, we shall see how it represents the intermolecular repulsion energy.

4. Fitting procedure and results

In order to test the ability of the overlap model to predict the intermolecular repulsion, we need to have an *ab initio* repulsion potential $V_{\text{rep}}(R, \Omega)$, which has been calculated at a wide range of relative molecular orientations Ω , as well as a model repulsion potential $V^{\text{model}}(R, \Omega)$, which has been derived either from the atomic overlap $S_{\rho, \text{atomic}}(R, \Omega)$, or the distributed overlap $S_{\rho, \text{dist}}(R, \Omega)$. $S_{\rho, \text{atomic}}(R, \Omega)$ was not used for $(N_2)_2$ because of the inaccuracy of the atomic charge density.

We have used an *ab initio* repulsion potential which was obtained from Hayes-Stone perturbation theory [22]. The first-order Hayes-Stone energy is a sum of the exchange-repulsion and electrostatic energies. At long range, the electrostatic energy

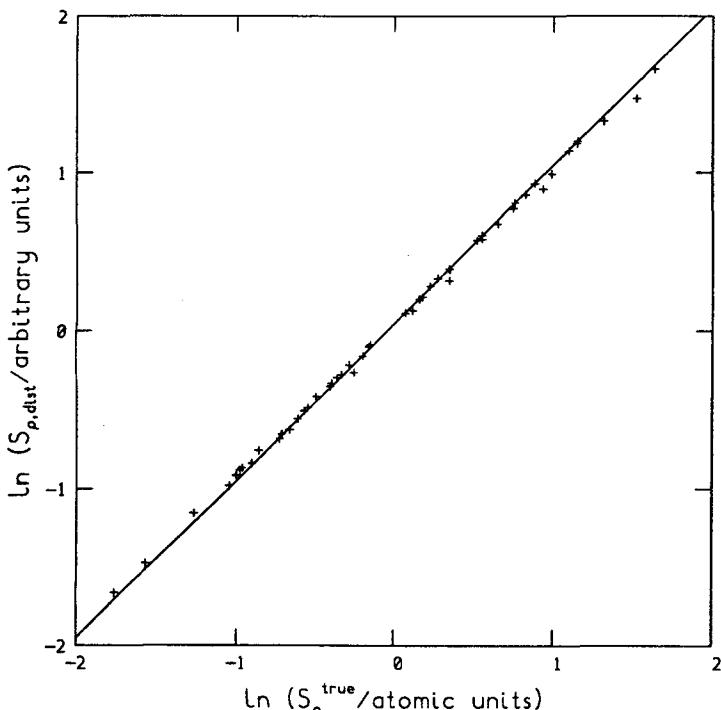


Figure 4. Comparison of the true and distributed overlap for the fluorine-fluorine interaction. The line has a gradient of 1. The values of $S_{\rho, \text{dist}}$ have been arbitrarily scaled, whereas S_{ρ}^{true} is in atomic units

can be obtained from a distributed multipole expansion [9, 23] of the electrostatic potential around the molecule. At short range, the electrostatic energy includes a penetration contribution due to the overlap of the charge clouds. For the purposes of developing model potentials, it is convenient to include this penetration as part of the repulsion energy V_{rep} . Hence, we define V_{rep} to be the difference between the first-order Hayes-Stone energy and the DMA electrostatic energy, the latter being obtained using a 3-site expansion up to hexadecapole. The model first-order interaction energy is now given by a sum of only two terms: the familiar DMA electrostatic energy, and our overlap model ‘repulsion’ energy. The Hayes-Stone first-order energy was calculated using Stone’s intermolecular perturbation theory program, which is an extension to the CADPAC [18] suite of programs. The energies were obtained at many different molecular orientations around the region of the repulsive wall which would be sampled in simulations of condensed phases. The distributed multipoles were obtained using CADPAC, and the electrostatic energy was derived from these using Stone’s ORIENT program [24]. Tong’s results [25] were used for chlorine and nitrogen. For fluorine, we used the same relative molecular orientations as Tong did for nitrogen, but adjusted the intermolecular separations so that V_{rep} lay approximately in the range 0.2×10^{-3} to $4 \times 10^{-3} E_h$. Table 4 summarizes the different repulsion surfaces.

The basis sets we have used certainly do not give an ideal representation of the charge density, especially since electron correlation has not been included. However, using the same basis set for all the calculations on a given molecule should mean that errors due to basis set deficiencies are eliminated. In any case, the anisotropy parameters in the overlap do not vary significantly with basis set, as is shown in figure 5 for fluorine. Gellert [16] has shown that the dependence of the radial part of the overlap may be more significant. If a highly accurate monomer wavefunction is required, correlated *ab initio* techniques may be used. This does not present any problems for the overlap model, since it is almost as easy to obtain the monomer charge density from a CI wavefunction as it is from an SCF wavefunction.

In order to derive a model repulsion potential, we used the basic assumption $V_{\text{rep}} = K S_p^x$, which we have already seen (in § 1) works very well for interactions

Table 4. Details of the V_{rep} surfaces used to test the model repulsion potentials.

	F_2	N_2	Cl_2
N (number of points calculated)	56	56	55
Number of distinct relative orientations	36	36	35
Minimum $V_{\text{rep}}/10^{-6} E_h$	221	393	497
Maximum $V_{\text{rep}}/10^{-6} E_h$	4064	8150	8648
$\langle V_{\text{rep}} \rangle /10^{-6} E_h$ †	1668	3473	3676
Boltzmann-weighted $\langle V_{\text{rep}} \rangle /10^{-6} E_h$ ‡	1008	1265	1325
Range of closest atom–atom contacts/ a_0	4.32–5.58	5.13–6.66	5.64–7.77
Atomic van der Waals diameter R_{vdw}/a_0	5.1	5.7	6.8

$$\dagger \langle V_{\text{rep}} \rangle = \left[\sum_{i=1}^N V_{\text{rep}}^2(i)/N \right]^{1/2}.$$

$$\ddagger \text{ Boltzmann-weighted } \langle V_{\text{rep}} \rangle = \left\{ \frac{\sum_{i=1}^N V_{\text{rep}}^2(i) \exp [-V_{\text{rep}}(i)/10^{-3} E_h]}{\sum_{i=1}^N \exp [-V_{\text{rep}}(i)/10^{-3} E_h]} \right\}^{1/2}.$$

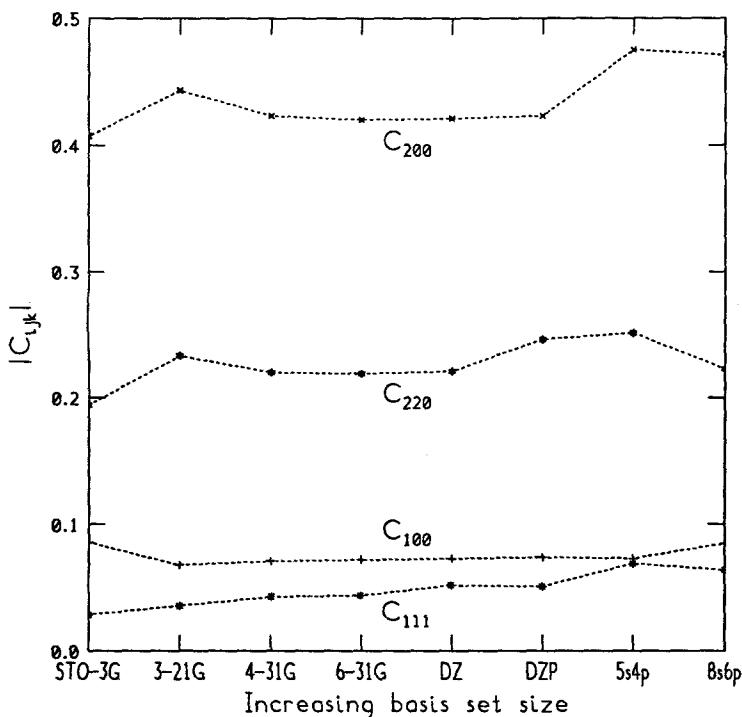


Figure 5. Basis set dependence of the atomic anisotropy parameters. The four largest anisotropy parameters for the atomic overlap of $(F_2)_2$ are plotted as a function of basis set. They were calculated at an atomic separation of $5.1 a_0$.

between spherical systems. This assumption is tested explicitly for $(F_2)_2$, $(Cl_2)_2$ and $(N_2)_2$ in figure 6, using $S_{\rho, \text{dist}}$ instead of S_ρ , since the latter is too expensive to calculate. The agreement is excellent for fluorine, and very good for nitrogen and chlorine. The gradients of the graphs suggest that suitable values of x should be 0.866 for fluorine, 0.845 for nitrogen and 0.791 for chlorine. However, we first tried to simplify this by assuming that repulsion is proportional to overlap, so $x = 1$. The resulting functional forms are exactly as given in equation (12), using the parameters from table 3, and may be summarized as follows:

$$V_{OM1}(R, \Omega) = K_1 S_{\rho, \text{dist}}(R, \Omega) \quad (16)$$

and

$$V_{OMA1}(R, \Omega) = K_{A1} S_{\rho, \text{atomic}}(R, \Omega). \quad (16a)$$

Only the proportionality parameter (K_1 or K_{A1}) needs to be fitted. The fitting procedure used was a straightforward non-linear optimization (NAG [26] routine E04HFF), which minimized the Boltzmann-weighted RMS error

$$\sigma_B = \left\{ \frac{\sum_{i=1}^N [V_{\text{rep}}(i) - V^{\text{model}}(i)]^2 \exp [-V_{\text{rep}}(i)/10^{-3} E_h]}{\sum_{i=1}^N \exp [-V_{\text{rep}}(i)/10^{-3} E_h]} \right\}^{1/2}. \quad (17)$$

The weighting energy $10^{-3} E_h$ corresponds to a temperature of 316 K, and was

introduced to reduce the importance of fitting the points high up the repulsive wall, and to facilitate comparison with Tong's work [25]. The results are shown in table 5. They compare very favourably with a 2-parameter isotropic atom-atom model

$$V_{\text{iso}} = \sum_{a,b} \exp [-\alpha(R_{ab} - \rho)] \quad (18)$$

for fluorine and chlorine, but for nitrogen the fit is worse.

Clearly, this simple approximation may be inaccurate, so we return to the more accurate relationship $V_{\text{rep}} = KS_{\rho}^x$. We have already derived a set of plausible values for x , but we do not attach too much importance to them, since they are likely to be basis set dependent [16], and we have found that a small variation in x does not affect the accuracy significantly. We have therefore used standardised values of $x = 0.85$ for fluorine and nitrogen, and $x = 0.80$ for chlorine. This gives our second pair of model potentials

$$V_{\text{OM2}}(R, \Omega) = K_2[S_{\rho, \text{dist}}(R, \Omega)]^x \quad (19)$$

and

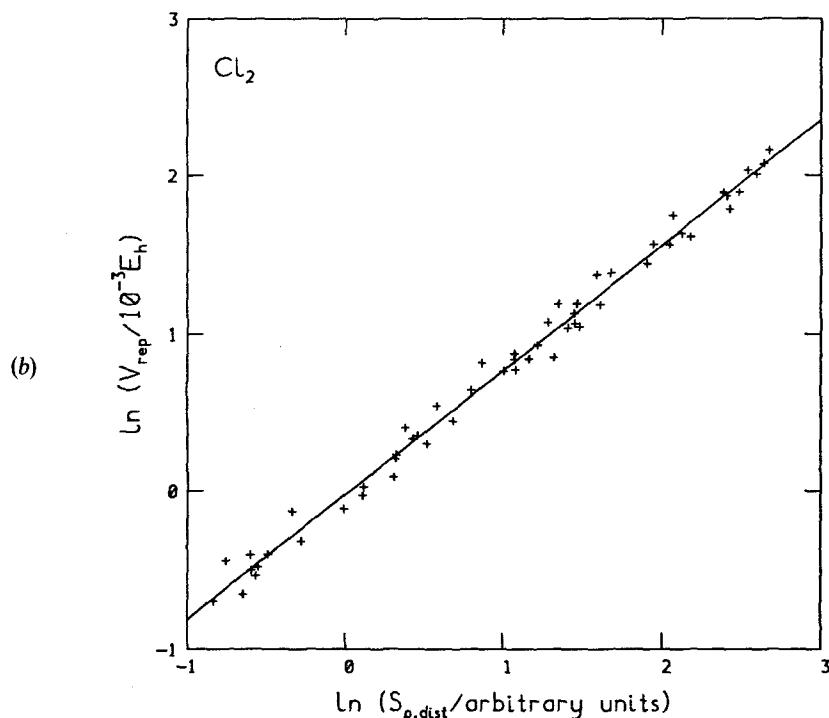
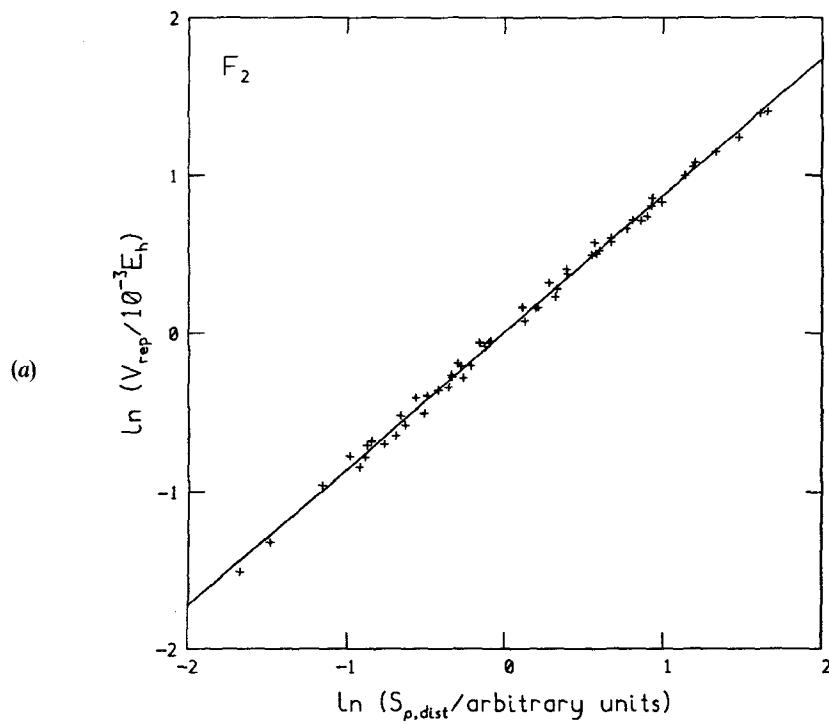
$$V_{\text{OMA2}}(R, \Omega) = K_{A2}[S_{\rho, \text{atomic}}(R, \Omega)]^x. \quad (19a)$$

The results from fitting the remaining parameter K are shown in table 5; there is a consistent reduction in σ_B by a factor of two. The fluorine and chlorine results are

Table 5. Comparison of the overlap model with the isotropic exponential model.

Model	Number of sites	Number of parameters	Values of parameters	Error $\sigma_B / 10^{-6} E_h$
			Fluorine	
ISO (isotropic)	2	2	$\rho = 4.969 a_0$ $\alpha = 2.174 a_0^{-1}$	197.6
OM1 ($K_1 S_{\rho, \text{dist}}$)	3	1	$K_1 = 2739 E_h$	101.9
OMA1 ($K_{A1} S_{\rho, \text{atomic}}$)	2	1	$K_{A1} = 2598 E_h$	93.9
OM2 ($K_2 [S_{\rho, \text{dist}}]^{0.85}$)	3	1†	$K_2 = 315.7 E_h$	41.5
OMA2 ($K_{A2} [S_{\rho, \text{atomic}}]^{0.85}$)	2	1†	$K_{A2} = 301.5 E_h$	46.6
OM3 ($K_3 [S_{\rho, \text{dist}}]$)	3	1†	$K_3 = 308.2 E_h$	93.4
OMA3 ($K_{A3} [S_{\rho, \text{atomic}}]$)	2	1†	$K_{A3} = 286.4 E_h$	126.9
OM3II	3	2†	$K_3 = 365.6 E_h$ $C_{000}^{\text{at-bc}} = -0.0655$	29.6
			Chlorine	
ISO	2	2	$\rho = 6.672 a_0$ $\alpha = 1.492 a_0^{-1}$	372.5
OM1	3	1	$K_1 = 19180 E_h$	228.0
OMA1	2	1	$K_{A1} = 18080 E_h$	228.2
OM2 ($K_2 [S_{\rho, \text{dist}}]^{0.80}$)	3	1†	$K_2 = 772.2 E_h$	109.6
OMA2 ($K_{A2} [S_{\rho, \text{atomic}}]^{0.80}$)	2	1†	$K_{A2} = 741.8 E_h$	112.0
OM3	3	1†	$K_3 = 774.9 E_h$	175.8
OMA3	2	1†	$K_{A3} = 718.7 E_h$	232.9
OM3II	3	2†	$K_3 = 902.5 E_h$ $C_{000}^{\text{at-bc}} = -0.0702$	85.0
			Nitrogen	
ISO	2	2	$\rho = 5.851 a_0$ $\alpha = 2.094 a_0^{-1}$	120.4
OM1	3	1	$K_1 = 1134 E_h$	206.3
OM2 ($K_2 [S_{\rho, \text{dist}}]^{0.85}$)	3	1†	$K_2 = 159.1 E_h$	133.8
OM3	3	1†	$K_3 = 138.3 E_h$	70.7
OM3II	3	2†	$K_3 = 133.2 E_h$ $C_{000}^{\text{at-bc}} = 0.1936$	45.7

† The value of x was derived separately (see text).
A detailed description of the models used is given in § 4.



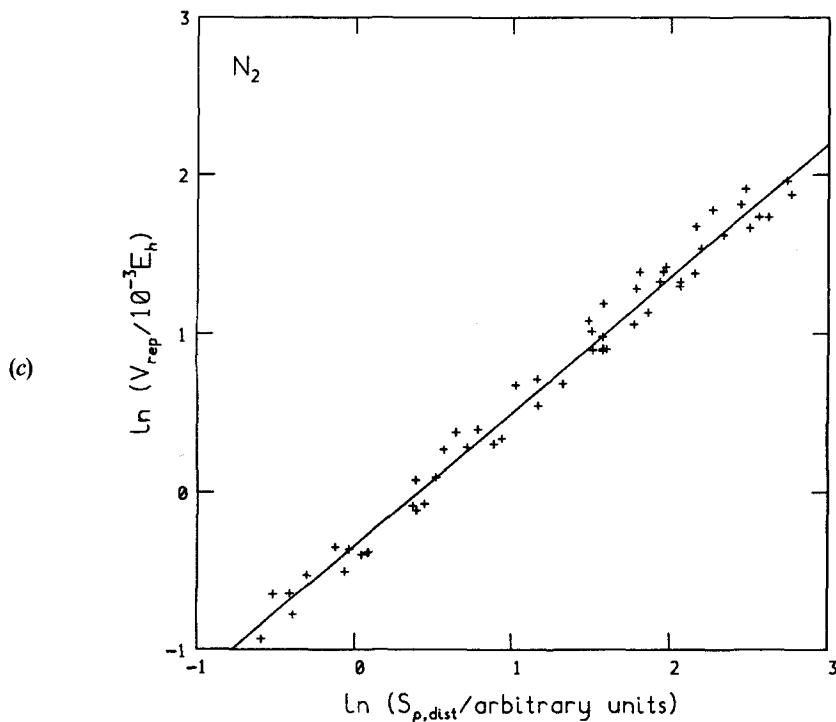


Figure 6. Comparison of *ab initio* repulsion and distributed overlap for the dimers of the homonuclear diatomics. The best line for $(F_2)_2$ (a) has a correlation coefficient of 0.998, for $(Cl_2)_2$ (b) it is 0.995, and for $(N_2)_2$ (c), it is 0.992.

now four times as accurate as the isotopic results, and the nitrogen results are similar, having had one fewer parameter fitted. We may conclude from this that the overlap model is excellent for singly bonded species, that the atomic overlap model is a simpler alternative (involving no atom–bond or bond–bond terms) which is just as accurate, and that the 2-site isotropic model is completely unsuitable.

The functional form of V_{OM2} and V_{OMA2} may be awkward to use, so we have introduced a third model which is based on the idea that raising the overlap to the power x , where $x < 1$, will mainly have the effect of flattening the exponential decay of the potential with distance. A similar effect is produced by multiplying all the α^{ab} by x , which gives a potential identical in form to V_{OM1} and V_{OMA1} . It may be more realistic than V_{OM2} , in that terms arising from individual site–site interactions are added together to give the potential energy, as for V_{OM1} , whereas for V_{OM2} they are added and then raised to a power, which is contrary to physical intuition, and is not in the form of equation (1). For the atomic model, only the exponents α^{ab} need readjusting, to give new exponents α_{OMA3}^{ab} , but for the distributed model it is also necessary to readjust the C_{000}^{ab} , in order to keep their relative values at the van der Waals diameter constant. The modified parameters for the distributed model are shown in table 3, where the same values of x were used as for V_{OM2} . This pair of models may be summarized by

$$V_{OM3}(R, \Omega) = K_3[S_{\rho, \text{dist}}(R, \Omega)]^x \quad (20)$$

and

$$V_{\text{OMA}3}(R, \Omega) = K_{A3}[S_{p, \text{atomic}}(R, \Omega)]', \quad (20a)$$

where

$$\alpha_{\text{OM(A)}3}^{ab}(\text{in } S_p) = x\alpha^{ab}(\text{in } S_p) \quad (21)$$

and

$$C_{000, \text{OM(A)}3}^{ab}(\text{in } S_p) = C_{000}^{ab}(\text{in } S_p) \exp [(1 - x)R_{\text{vdw}}(\alpha^{\text{at-at}} - \alpha^{ab})]. \quad (22)$$

Table 5 summarizes the results obtained by fitting the proportionality constant K . For fluorine and chlorine, the accuracy is similar to $V_{\text{OM}1}$ and $V_{\text{OM}1}$, but for nitrogen there is a dramatic improvement in the fit. For all three molecules, $V_{\text{OM}3}$ is approximately twice as accurate as the isotropic atom–atom model. This is, therefore, the most consistent model, and it would be our first choice for applying to a new system. However, it is clearly also worth trying $V_{\text{OM}2}$, or $V_{\text{OMA}2}$ for singly bonded systems.

Since we have a relatively large amount of *ab initio* data available, which samples many different orientations, it is possible to fit more parameters. We first fitted two parameters in our first choice model $V_{\text{OM}3}$, in order to make a comparison with the isotropic model more realistic. We decided to fit $C_{000}^{\text{at-bc}}$, in order to test whether the overlap model gives the correct importance to the atom–bond terms. The results are shown under OM3II in table 5. They are convincing evidence of the success of our model, and show the serious limitations of the isotropic model, which is nearly an order of magnitude worse. The accuracy of the fits for fluorine and nitrogen is quite outstanding, and the chlorine fit is also extremely good. We also tried optimizing x in the OM3 model (not shown). For nitrogen, this made very little difference, but an improvement was observed for fluorine and chlorine. However, this was not as great an improvement as that obtained by fitting $C_{000}^{\text{at-bc}}$. The conclusions are that the relative importance of the atom–atom and atom–bond terms is probably the source of the problem with the overlap model, particularly for $(\text{N}_2)_2$, and that the value of x is not crucial, thus justifying the use of a separately derived, not fully optimised value, which may only depend on the row of the periodic table.

The anisotropy of repulsion is very well represented for $(\text{F}_2)_2$ and $(\text{Cl}_2)_2$. We can confirm this by fitting the largest atom–atom anisotropy parameter $C_{200}^{\text{at-at}}$, as well as K_3 and $C_{000}^{\text{at-bc}}$ in model OM3II. This changed the value of $C_{200}^{\text{at-at}}$ from -0.404 to -0.412 for $(\text{F}_2)_2$ and from -0.546 to -0.536 for $(\text{Cl}_2)_2$. Similarly, optimizing $C_{220}^{\text{at-at}}$ changed its value from 0.199 to 0.201 for $(\text{F}_2)_2$ and from 0.363 to 0.394 for $(\text{Cl}_2)_2$. However, the values of these parameters are strongly coupled, and when we fit both the largest anisotropy parameters simultaneously, together with K_3 and $C_{000}^{\text{at-bc}}$, the change in the anisotropy parameters is significant for all three molecules, as shown in table 6 under OM3III. The accuracy of these fits is excellent: the rms percentage errors are 1.34 per cent for fluorine, and 4.15 per cent for chlorine. We have also tried fitting four parameters in the atomic model, but the accuracy was not nearly as good, which suggests that atom–bond terms are required to give high accuracy.

We have also included in table 6 the results from fitting one state-of-the-art empirical model potential; this is the modified exponential (ME) atom–atom form

$$V_{\text{ME}} = \sum_{ab} \exp \{-\alpha_{ab}[R_{ab} - \rho_{ab}(\Omega_{ab})]\}. \quad (23)$$

Table 6. Comparison of the overlap model with modern empirical models.

Model	Number of sites	Number of parameters	Values of parameters	Error $\sigma_B / 10^{-6} E_h$
Fluorine				
TP			No test-particle results are available for fluorine.	
ME	2	4	$\alpha = 2.537 a_0^{-1}$ $\rho_0 = 5.108 a_0$ $\rho_1 = -0.134 a_0$ $\rho_2 = -0.281 a_0$	18.9
OM3II	3	2†	$K_3 = 365.6 E_h$ $C_{000}^{at-bc} = -0.0655$	29.6
OM3III	3	4†	$K_3 = 394.5 E_h$ $C_{200}^{at-at} = C_{020}^{at-at} = -0.461$ $C_{220}^{at-at} = 0.292$ $C_{000}^{at-bc} = -0.0728$	11.9
Chlorine				
TP	2	5	See [25]	221.9‡
ME	2	4	$\alpha = 1.907 a_0^{-1}$ $\rho_0 = 7.009 a_0$ $\rho_1 = -0.266 a_0$ $\rho_2 = -0.532 a_0$	59.8
OM3II	3	2†	$K_3 = 902.5 E_h$ $C_{000}^{at-bc} = -0.0702$	85.0
OM3III	3	4†	$K_3 = 952.9 E_h$ $C_{200}^{at-at} = C_{020}^{at-at} = -0.622$ $C_{220}^{at-at} = 0.526$ $C_{000}^{at-bc} = -0.0725$	46.6
Nitrogen				
TP	2	4	See [25]	42.8§
ME	2	4	$\alpha = 2.074 a_0^{-1}$ $\rho_0 = 5.792 a_0$ $\rho_1 = 0.111 a_0$ $\rho_2 = 0.193 a_0$	24.0
OM3II	3	2†	$K_3 = 133.2 E_h$ $C_{000}^{at-bc} = 0.1936$	45.7
OM3III	3	4†	$K_3 = 100.4 E_h$ $C_{100}^{at-at} = -C_{010}^{at-at} = -0.802$ $C_{110}^{at-at} = -0.482$ $C_{000}^{at-bc} = 0.3783$	36.0

† These use the values $x = 0.85$ for F_2 and N_2 , $x = 0.80$ for Cl_2 which were derived separately, as explained in the text.

‡ Best result, from 30 reported.

§ Best result, from 25 reported.

Our best 2-parameter fits have been included, for comparison. Section 4 contains a more detailed description of the models used.

Note that $10^{-6} E_h = 2.63 \text{ J mol}^{-1} = 0.316 \text{ K} = 0.219 \text{ cm}^{-1} = 6.58 \text{ GHz}$.

As it stands, this model is only suitable for interactions between homonuclear diatomics, since about four parameters must be fitted for each distinct site-site interaction, most of these parameters being used to describe the empirically chosen form for the anisotropy of ρ_{ab} . When we choose a standard form for the anisotropy

$$\rho_{ab}(\Omega_{ab}) = \sum_{k=0}^2 \rho_k [(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^k + (-\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^k], \quad (24)$$

this model gives inferior results to the overlap model OM3III for the same number of parameters, for $(F_2)_2$ and $(Cl_2)_2$. For $(N_2)_2$ it is rather better than the overlap model.

A test-particle (TP) model has been used by Tong [25] to predict the parameters for the $X-X$ interaction using data from the $He-X$ and $He-He$ interactions; this reduces the number of fitted parameters to four for each distinct site, but the model will still only be suitable for small molecules (up to, say, three heavy atoms), because of the expense of computing all the $ab initio$ points required. Also, the test particle model is incapable of predicting terms which are not described by the S -functions [27] S_{011} or S_{101} . However, some of these additional terms do appear in the full expression for the overlap, for example S_{112} and S_{222} . They do not appear in our

assumed expression for $\rho(\Omega)$, but they could be included in an empirical ME model. However, their inclusion would involve fitting more parameters. The test particle model works well for nitrogen, but for chlorine it is very poor (table 6). In summary, the distributed overlap model gives exceptionally good agreement with *ab initio* results when four parameters are fitted, and in the case of $(F_2)_2$ and $(Cl_2)_2$, the agreement is considerably better than that obtained using the best currently available repulsion potentials for small molecules. Since the distributed overlap model has a reasonable theoretical foundation, it is readily extended to large molecules whose size is only limited by the necessity of obtaining an *ab initio* monomer wavefunction, whereas the same is not true for the empirical modified exponential and test-particle models, as it is an impossible task to obtain large numbers of parameters with no theoretical guidance.

5. Discussion

One of the main problems in developing accurate model intermolecular potentials is that of determining a realistic model for the repulsion potential. It is necessary to determine the form of the anisotropy in the repulsion around each atom, which depends on the valence electron distribution, as well as the corresponding parameters. Simulations of condensed phases are sensitive to the shape of the repulsive wall in the model potential, and yet are insufficient to determine the repulsion potential unambiguously. This paper shows that consideration of the anisotropy of the overlap of the undistorted charge distributions provides valuable guidance as to the form of the anisotropy of repulsion, and that analysis of the monomer wavefunctions can provide useful quantitative estimates of the anisotropy parameters.

The model is particularly successful for F_2 and Cl_2 , where there is little bonding density. Restricting the analysis to the terms in the monomer charge density which involve pairs of orbitals on the same atom provides a model with only one parameter to be determined from other sources, which is markedly superior to the isotropic atom–atom potential. Optimizing an extra parameter or so, against either suitable experimental data or an *ab initio* repulsion surface, gives a model of considerable accuracy, which would be hard to achieve without the overlap model's guidance as to which anisotropy parameters are most important.

The bonding electron density makes little contribution to the overlap for $(F_2)_2$ and $(Cl_2)_2$, which is shown by the relative unimportance of those terms in the parameters obtained from the distributed charge density model. This is in marked contrast to the results for N_2 , where the bonding electron density is significant (table 3). The overlap model is less satisfactory for N_2 than for F_2 or Cl_2 , particularly relative to the isotropic atom–atom model, which is a much better starting point for N_2 than for the other systems. This will be due to the basic relationship $V_{rep} \propto S_\rho^x$ being the least accurate for N_2 (figure 6). The marked improvement in the predicted V_{rep} when the relative importance of the atom–atom and atom–bond centre terms is treated as a parameter suggests that the root of the problem may be a difference in the repulsion/overlap relationship for atomic versus bonding electron density. This is not unreasonable, since internuclear repulsion contributes to the repulsion between atomic sites, but not to interactions involving a site at a bond centre.

One problem with the empirical basis of the overlap model (i.e. the assumption $V_{rep} \propto S_\rho^x$), that if two large systems overlap at two widely separated points, the total

model repulsion energy is not a sum of two contributions, one from each point, unless $x = 1$. This is another justification for our OM3 model, which treats each interaction separately and adjusts the radial dependence of each separately, before adding them together. When the overlap model is applied to larger systems, this factor which must be taken into consideration, since the effects of non-additivity will then be more noticeable. Another possibility, which has the undesirable side-effect of complicating the potential form, is to raise each of the individual site-site overlaps to the power x before adding them. However, all of these extensions cannot overcome the fact that any assumption that the repulsion is related to the overlap is empirical, and any attempts to achieve higher accuracy will be limited by the accuracy of this assumption. Nyeland and Toennies [28] have investigated the alternative relationship $V_{\text{rep}} \approx KS_p/R^2$ for the repulsion between rare gas atoms, but Gellert [16] has shown that this does not, in general, improve the accuracy. In general, the basic assumption $V_{\text{rep}} \propto S_p^x$ is very good. The anisotropy parameters are especially well represented, and it is primarily in the radial part of the model potential that inaccuracies arise. The best way to achieve higher accuracy is probably to fit more parameters, where sufficient data are available.

A particularly useful feature of the overlap model of the repulsion is that we cannot foresee any problems in using it to determine an atom-atom model repulsion potential for larger molecules, provided that the effects of the bonding density can be ignored. The atomic overlap models OMA2 and OMA3 should be applicable to a wide range of organic molecules, and the distributed charge density method can be used to determine which functional groups may require a more elaborate method. The guidance offered by the overlap model as to the form of the anisotropy will be more important as the molecular symmetry is lowered, and the form of the anisotropy therefore becomes more complicated. It will be necessary to expand the anisotropy of the overlap model potential in terms of S -functions [27], which are a complete orthogonal set of orientation functions, since an unambiguous expansion in terms of direction cosines is not possible for non-linear molecules. Empirical fitting of both the form and the parameters of the repulsion, as in the modified exponential model, would be quite unsuitable for systems of low symmetry, since all possible anisotropic terms would have to be included, with no guide as to their importance.

The rapid improvement in computer power means that it is now possible to perform molecular dynamics simulations with realistic anisotropic model pair potentials [5]. We could envisage a hypothetical model intermolecular pair potential given by

$$V^{\text{model}} = V_{\text{rep}} + V_{\text{elec}} + V_{\text{disp}}, \quad (25)$$

where perturbation theory can provide guidance on the form of the potential at long range, and, in principle, the long-range parameters can be obtained from *ab initio* calculations on the monomer. What is required is a distributed model, with sites on the atoms, and possibly also the bond centres, depending on the electron distribution. The electrostatic term should be modelled by a distributed multipole model such as a DMA [9]. At short range, the potential needs to model the exchange-repulsion and the modification of the electrostatic and dispersion terms by the overlap of the charge distributions. We have shown that the analysis of the monomer wavefunction can give a worthwhile model for V_{rep} (the exchange-repulsion plus electrostatic penetration), and it seems likely that the required empiri-

cal fitting of one or two parameters will at least partially absorb the dispersion damping and other errors. Thus, the analysis of a high quality monomer wavefunction could provide a firm basis for an anisotropic site-site potential, which represents the intermolecular effects of non-spherical features in the molecular charge density, leaving only a few parameters to be determined by experimental data.

There is still much work to be done before better model intermolecular potentials for even the smallest polyatomics reach the accuracy that has been achieved for argon [29]. The form of the best argon potentials, as well as the history of intermolecular potentials for polyatomics [30], suggests that this will require a complex functional form with dozens of parameters. We believe that the overlap model for the repulsion, by providing guidance on the form of site-site anisotropy and estimates of potential energy parameters, will make a significant contribution towards more realistic model intermolecular potentials for the computer simulation of condensed phases and the interpretation of spectroscopic data.

R.J.W. is grateful to the SERC for a research studentship. S.L.P. thanks the Royal Society for support in the form of a 1983 University Research Fellowship.

Appendix

An exact expression for the charge density of an SCF ab initio wavefunction

The charge density is defined by equations (2) and (3), which are repeated here for convenience.

$$\rho_A(\mathbf{r}) = \sum_{i,j} \phi_i(\mathbf{r})\phi_j(\mathbf{r})M_{ij}, \quad (2)$$

$$\phi_i(\mathbf{r}) = \sum_u a_{iu} \eta_u(\mathbf{r}) = \sum_u a_{iu} (x - x_i)^{\lambda_u} (y - y_i)^{\mu_u} (z - z_i)^{\nu_u} \exp[-\alpha_u(\mathbf{r} - \mathbf{r}_i)^2]. \quad (3)$$

Combining these two equations gives

$$\begin{aligned} \rho_A(\mathbf{r}) = & \sum_{ijuv} a_{iu} a_{jv} M_{ij} (x - x_i)^{\lambda_u} (x - x_j)^{\lambda_v} (y - y_i)^{\mu_u} (y - y_j)^{\mu_v} (z - z_i)^{\nu_u} (z - z_j)^{\nu_v} \\ & \times \exp[-\alpha_u(\mathbf{r} - \mathbf{r}_i)^2] \exp[-\alpha_v(\mathbf{r} - \mathbf{r}_j)^2]. \end{aligned} \quad (A 1)$$

Each product of two gaussians in this sum may be re-expressed as a single gaussian centred at $\mathbf{r}_n = (x_n, y_n, z_n) = (\alpha_u \mathbf{r}_i + \alpha_v \mathbf{r}_j)/(\alpha_u + \alpha_v)$, and it is easy to verify that

$$\begin{aligned} \rho_A(\mathbf{r}) = & \sum_{ijuv} a_{iu} a_{jv} M_{ij} (x - x_n + x_n - x_i)^{\lambda_u} (x - x_n + x_n - x_j)^{\lambda_v} \\ & \times (y - y_n + y_n - y_i)^{\mu_u} (y - y_n + y_n - y_j)^{\mu_v} \\ & \times (z - z_n + z_n - z_i)^{\nu_u} (z - z_n + z_n - z_j)^{\nu_v} \\ & \times \exp[-\alpha_u \alpha_v (\mathbf{r}_i - \mathbf{r}_j)^2 / (\alpha_u + \alpha_v)] \times \exp[-(\alpha_u + \alpha_v)(\mathbf{r} - \mathbf{r}_n)^2]. \end{aligned} \quad (A 2)$$

The pre-exponential terms are now expanded in binomial series to give

$$\begin{aligned} \rho_A(\mathbf{r}) = & \sum_{ijuv} a_{iu} a_{jv} M_{ij} \exp [-\alpha_u \alpha_v (\mathbf{r}_i - \mathbf{r}_j)^2 / (\alpha_u + \alpha_v)] \\ & \times \sum_{\lambda_1, \lambda_2=0}^{\lambda_u, \lambda_v} \sum_{\mu_1, \mu_2=0}^{\mu_u, \mu_v} \sum_{v_1, v_2=0}^{v_u, v_v} (x_n - x_i)^{\lambda_u - \lambda_1} (x_n - x_j)^{\lambda_v - \lambda_2} (y_n - y_i)^{\mu_u - \mu_1} (y_n - y_j)^{\mu_v - \mu_2} \\ & \times (z_n - z_i)^{v_u - v_1} (z_n - z_j)^{v_v - v_2} \binom{\lambda_u}{\lambda_1} \binom{\lambda_v}{\lambda_2} \binom{\mu_u}{\mu_1} \binom{\mu_v}{\mu_2} \binom{v_u}{v_1} \binom{v_v}{v_2} \\ & \times (x - x_n)^{\lambda_1 + \lambda_2} (y - y_n)^{\mu_1 + \mu_2} (z - z_n)^{v_1 + v_2} \exp [-(\alpha_u + \alpha_v)(\mathbf{r} - \mathbf{r}_n)^2]. \end{aligned} \quad (A 3)$$

This has the required form of equation (4), where the sum over i, j, u and v has been re-expressed in terms of a sum over n .

The explicit form of overlaps between gaussians up to rank 2 which comprise the charge distributions of linear molecules

From equation (A 3), it is clear that gaussians of rank 2 or less are the only ones to appear in the atomic charge distributions of wavefunctions calculated using sp basis sets. If the molecules are also linear, then the charge distributions just consist of s , z , x^2 , y^2 and z^2 gaussians, and the x^2 , and y^2 gaussians always occur as a pair:

$$s_{x^2+y^2} = k_n [(x - x_n)^2 + (y - y_n)^2] \exp [-\alpha_n(\mathbf{r} - \mathbf{r}_n)^2], \quad (A 4)$$

because of the symmetry properties of the density matrix elements. In equations (A 5) to (A 14) we shall give the analytical form of overlaps between each pair of gaussian types as a function of separation R and relative orientation Ω . This information is all that is needed to implement the atomic overlap model for linear molecules with sp basis sets.

The overlap of two s orbitals gives an isotropic contribution

$$\begin{aligned} s_\rho(R, \Omega; \eta_s, \eta_s) &= k_a k_b \exp \left[\frac{-\alpha_a \alpha_b}{\alpha_a + \alpha_b} R^2 \right] \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} \\ &= P \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)}. \end{aligned} \quad (A 5)$$

All further expressions contain the same factor P . The remaining overlaps are expressed in terms of R and the direction cosines $\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}$, $\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}$ and $\hat{\mathbf{z}}_A \cdot \mathbf{z}_B \cdot \mathbf{R} = R \hat{\mathbf{R}}$ is a vector from the centre of one gaussian, which has exponent α_a , to the centre of the other gaussian, which has exponent α_b .

$$\begin{aligned} s_\rho(R, \Omega; \eta_s, \eta_z) &= P \left\{ \frac{-\alpha_a R}{\alpha_a + \alpha_b} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}) \right\}, \\ s_\rho(R, \Omega; \eta_s, \eta_{z^2}) &= P \left\{ \frac{1}{2} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} + \left(\frac{\alpha_a R}{\alpha_a + \alpha_b} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2 \right\}, \end{aligned} \quad (A 7)$$

$$s_\rho(R, \Omega; \eta_s, \eta_{x^2+y^2}) = P \left\{ \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} + \left(\frac{\alpha_a R}{\alpha_a + \alpha_b} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} [1 - (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2] \right\}, \quad (A 8)$$

$$s_\rho(R, \Omega; \eta_z, \eta_{z^2}) = P \left\{ \frac{1}{2} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B) - \frac{\alpha_a \alpha_b R^2}{(\alpha_a + \alpha_b)^2} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} \times (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}}) \right\}, \quad (\text{A } 9)$$

$$s_\rho(R, \Omega; \eta_z, \eta_{z^2}) = P \left\{ \frac{1}{2} \frac{\alpha_b R}{\alpha_a + \alpha_b} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}) - \frac{\alpha_a R}{\alpha_a + \alpha_b} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B) + \frac{\alpha_a^2 \alpha_b R^3}{(\alpha_a + \alpha_b)^3} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2 \right\}, \quad (\text{A } 10)$$

$$s_\rho(R, \Omega; \eta_z, \eta_{x^2+y^2}) = P \left\{ \frac{\alpha_b R}{\alpha_a + \alpha_b} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}) + \frac{\alpha_a R}{\alpha_a + \alpha_b} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} [(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B) - (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})] + \frac{\alpha_a^2 \alpha_b R^3}{(\alpha_a + \alpha_b)^3} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} [(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}}) - (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2] \right\}, \quad (\text{A } 11)$$

$$s_\rho(R, \Omega; \eta_{z^2}, \eta_{z^2}) = P \left\{ \frac{1}{2} \left(\frac{\alpha_b R}{\alpha_a + \alpha_b} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2 + \frac{1}{2} \left(\frac{\alpha_a R}{\alpha_a + \alpha_b} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2 + \left(\frac{\alpha_a \alpha_b R^2}{(\alpha_a + \alpha_b)^2} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2 (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2 - 2 \frac{\alpha_a \alpha_b R^2}{(\alpha_a + \alpha_b)^2} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B) + \frac{1}{4} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^7} \right)} [2(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)^2 + 1] \right\}, \quad (\text{A } 12)$$

$$s_\rho(R, \Omega; \eta_{z^2}, \eta_{x^2+y^2}) = P \left\{ \left(\frac{\alpha_b R}{\alpha_a + \alpha_b} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2 + \frac{1}{2} \left(\frac{\alpha_a R}{\alpha_a + \alpha_b} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} [1 - (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2] + \left(\frac{\alpha_a \alpha_b R^2}{(\alpha_a + \alpha_b)^2} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} [(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2 - (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2 (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2] + 2 \frac{\alpha_a \alpha_b R^2}{(\alpha_a + \alpha_b)^2} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} \times [(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B) - (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2] + \frac{1}{2} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^7} \right)} [2 - (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)^2] \right\}, \quad (\text{A } 13)$$

$$\begin{aligned}
 s_p(R, \Omega; \eta_{x^2+y^2}, \eta_{x^2+y^2}) = & P \left\{ \left(\frac{\alpha_b R}{\alpha_a + \alpha_b} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} [1 - (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2] \right. \\
 & + \left(\frac{\alpha_a R}{\alpha_a + \alpha_b} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} [1 - (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2] \\
 & + \left(\frac{\alpha_a \alpha_b R^2}{(\alpha_a + \alpha_b)^2} \right)^2 \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^3} \right)} \\
 & \times [1 - (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2 - (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2 + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2 (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2] \\
 & + 2 \frac{\alpha_a \alpha_b R^2}{(\alpha_a + \alpha_b)^2} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^5} \right)} \\
 & \times [(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})^2 + (\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})^2 - (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B) - 1] \\
 & \left. + \frac{1}{2} \sqrt{\left(\frac{\pi^3}{(\alpha_a + \alpha_b)^7} \right)} [3 + (\hat{\mathbf{z}}_A \cdot \hat{\mathbf{z}}_B)^2] \right\}. \quad (\text{A } 14)
 \end{aligned}$$

The overlaps where the order of the orbitals is reversed can be derived from the above expressions by interchanging $\hat{\mathbf{z}}_A$ and $\hat{\mathbf{z}}_B$, α_a and α_b , and \mathbf{R} and $-\mathbf{R}$. For example, to obtain $s_p(R, \Omega; \eta_z, \eta_s)$ from $s_p(R, \Omega; \eta_s, \eta_z)$, $-\alpha_a R$ should be replaced by $-\alpha_b R$, and $(\hat{\mathbf{z}}_B \cdot \hat{\mathbf{R}})$ by $-(\hat{\mathbf{z}}_A \cdot \hat{\mathbf{R}})$.

References

- [1] IMPEY, R. W., and KLEIN, M. L., 1983, *Chem. Phys. Lett.*, **103**, 143.
- [2] KOBASHI, K., and KLEIN, M. L., 1980, *Molec. Phys.*, **41**, 679.
- [3] ANDREANI, C., CILLOCO, F., NENCINI, L., ROCCA, D., and SINCLAIR, R. N., 1985, *Molec. Phys.*, **55**, 887.
- [4] NYBURG, S. C., and FAERMAN, C. H., 1985, *Acta crystallogr. B*, **41**, 274.
- [5] RODGER, P. M., STONE, A. J., and TILDESLEY, D. J., 1987, *J. chem. Soc. Faraday Trans. II*, **83**, 1689.
- [6] BROBJER, J. T., and MURRELL, J. N., 1983, *Molec. Phys.*, **50**, 885.
- [7] WIBERG, K. B., and MURCKO, M. A., 1987, *J. comput. Chem.*, **8**, 1124.
- [8] PRICE, S. L., 1986, *Molec. Phys.*, **58**, 651.
- [9] STONE, A. J., 1981, *Chem. Phys. Lett.*, **83**, 233.
- [10] STONE, A. J., 1989, *Chem. Phys. Lett.*, **155**, 102, 111.
- [11] STONE, A. J., and TONG, C. S., 1989, *Chem. Phys.*, **137**, 121.
- [12] STONE, A. J., 1989, *Hydrogen Bonded Liquids*, (NATO ASI Summer School, Cargese).
- [13] STONE, A. J., and TONG, C. S. (in preparation).
- [14] KITA, S., NODA, K., and INOUYE, H., 1976, *J. chem. Phys.*, **64**, 3446.
- [15] KIM, Y. S., KIM, S. K., and LEE, W. D., 1981, *Chem. Phys. Lett.*, **80**, 574.
- [16] GELLERT, P., 1988, Ph.D. Thesis, University of Oxford.
- [17] BOYS, S. F., 1950, *Proc. R. Soc.*, **200**, 542.
- [18] AMOS, R. D., 1984, *CADPAC, the Cambridge Analytical Derivatives Package* (CCP1/84/4, Daresbury).
- [19] 1988, *Handbook of Chemistry and Physics*, Vol. 69 (C.R.C.).
- [20] See, for example; LONDON, F., 1937, *J. chem. Soc. Faraday Trans. II*, **33**, 8.
- [21] PRICE, S. L., 1987, *Molec. Phys.*, **62**, 45.
- [22] HAYES, I. C., and STONE, A. J., 1984, *Molec. Phys.*, **53**, 83.
- [23] ALDERTON, M., and STONE, A. J., 1985, *Molec. Phys.*, **56**, 1047.
- [24] PRICE, S. L., and STONE, A. J., 1987, *J. chem. Phys.*, **86**, 2859.
- [25] TONG, C. S., 1988, Ph.D. Thesis, University of Cambridge.
- [26] Numerical Algorithms Group, Oxford, England.
- [27] STONE, A. J., 1978, *Molec. Phys.*, **36**, 241.
- [28] NYELAND, C., and TOENNIES, J. P., 1986, *Chem. Phys. Lett.*, **127**, 172.
- [29] MAITLAND, G. C., RIGBY, M., SMITH, E. B., and WAKEHAM, W. A., 1981, *Intermolecular Forces* (Clarendon).
- [30] PRICE, S. L., 1988, *Molec. Simul.*, **1**, 135.