Distributed dispersion: A new approach

Greg J. Williams[a] and Anthony J. Stone[b]
University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Distributed polarizability and dispersion models can be constructed by analyzing the response, at each point in an array surrounding the molecule, to the field due to a point charge at each point of the array in turn, in an extension of the method described by Dehez *et al.* [Chem. Phys. Lett. **338**, 180 (2001)]. Anisotropic distributed polarizabilities and anisotropic distributed C_6 dispersion coefficients have been calculated for carbon dioxide, water, benzene and ethane. © *2003 American Institute of Physics.* [DOI: 10.1063/1.1594722]

I. INTRODUCTION

Modern intermolecular potential functions are commonly constructed from separate terms describing the contributions arising from intermolecular perturbation theory.[1] The main terms are repulsion, electrostatic, and dispersion, though induction, charge transfer, and other smaller terms may also contribute. For all but the smallest molecules it is usual to represent each of these terms in atom–atom form, using distributed multipoles for the electrostatic description and atom–atom terms, often of Born–Mayer form, for the repulsion. The dispersion terms are more difficult to deal with if high accuracy is required. For many small molecules, isotropically averaged C_6 values are available from the work of Meath and others,[2] who obtained them from dipole oscillator strength distributions, and it is possible in some cases to obtain information about the anisotropy of the C_6 coefficient.[3–5] Many conventional force field models use simple $-C_6R^6$ forms fitted to experimental data, but these ignore any orientation dependence. It is clear, however, even from the London expression for the dispersion coefficient,[1] that if the polarizability is anisotropic then the dispersion must be anisotropic too.

It is sometimes remarked that the use of an atom–atom model provides a way to describe anisotropy in the intermolecular interaction, even if the atom–atom terms are isotropic. It is true that such a description incorporates some anisotropy, especially in the case of repulsion, but even there it is well established that atomic anisotropy is essential to an accurate description. In the case of dispersion, the use of isotropic atom–atom C_6 dispersion coefficients leads directly to an isotropic molecular C_6. Any anisotropy arising from the use of the atom–atom form enters only in higher terms in the R^{-n} expansion. For an accurate description, as in the case of repulsion, we need to embrace the anisotropic atom–atom model.

For higher terms in the dispersion energy series, we are dependent entirely on *ab initio* calculations, but they usually give only the dispersion coefficients referred to a molecular origin, which may be satisfactory for small molecules but are rarely adequate for larger ones. It is possible to define distributed polarizabilities in a rigorous manner,[6] and to compute them reasonably efficiently.[7] Using such polarizabilities calculated at imaginary frequency it is possible to obtain a distributed dispersion description,[8] but it is quite cumbersome, involving in its most general form sums over pairs of atoms in each molecule.

We explore here an alternative approach, based on the ideas of Nakagawa and Kosugi,[9] Dehez *et al.*, and Celebi *et al.*[10,11] The latter computed the induction energy of a molecule in the field of a nearby point charge. By placing the point charge at many different positions, they were able to obtain detailed information about the induction energy as a function of its position. If a polarizability model is postulated for the molecule, involving whatever distributed-polarizability terms are considered appropriate, the parameters in the model can be fitted so as to reproduce the observed induction energies as closely as possible. A variety of polarizability models can be investigated, and in this way the terms needed to obtain an accurate polarizability description can be identified and their values estimated.

In the present work we extend this approach in two ways. First, we set up in a similar way a grid of points surrounding the molecule. We then consider the field of a point charge at each position in turn, and calculate the response (the induced potential) at every point on the grid, not just at the position of the perturbing charge. The complete set of responses for a large number of points (1000 or more in our present implementation) can be obtained in a single calculation using the CADPAC program.[12] This large dataset, or several such datasets, can then be fitted to a distributed-polarizability model. The CADPAC program also allows for the computation of the responses at a set of imaginary frequencies, and these can be fitted to the same model to give distributed polarizabilities at imaginary frequency, which in turn can be used to obtain distributed dispersion coefficients.

This approach has the fairly obvious advantage that it provides much more detailed information about the molecule's response, but there is also the more subtle advantage that the fitted functions are approximately orthogonal, which makes for a much more stable and well-conditioned fitting procedure, as we shall show below.

However, the main new feature is that it provides a route

to a distributed description of dispersion. Many methods have been proposed in the past for obtaining atomic polarizabilities for molecules, from the early work of LeFevre[13] and Applequist[14–16] to more recent studies,[17,18] but only the work of Hättig et al.[19] explored the time-dependent polarizabilities that provide the route to atomic dispersion coefficients. We believe that the present paper provides a more versatile and practical approach.

II. THEORY AND COMPUTATIONAL METHOD

A point charge at position P produces an electrostatic potential $V_P(r) = (4\pi\epsilon_0|r-P|)^{-1}$ at a point r. The change in electrostatic potential at some point Q is described in second-order perturbation theory by an off-diagonal polarizability, expressed in sum-over-states form as

$$\alpha_{PQ} = 2\sum_n \frac{\langle 0|\Sigma_i V_P(r_i)|n\rangle\langle n|\Sigma_j V_Q(r_j)|0\rangle}{W_n - W_0}, \qquad (1)$$

where i and j refer to all the particles, electrons, and nuclei, in the molecule. In a Born–Oppenheimer clamped-nucleus picture, however, the nuclear coordinates do not enter, and the sum is over electrons only. The matrix elements required have the form of nuclear-attraction integrals, and are easily evaluated using standard code. The perturbation calculation can be efficiently and accurately carried out using current density-functional theory[20] with a suitable functional, such as the widely-used B3LYP functional,[21] or the PBE0 hybrid functional,[22] which has been found to give good results for polarizabilities. This method has been implemented in the CADPAC program.[12] Given the basis-function matrix elements for each of N point-charge perturbations, CADPAC calculates the N^2 point-to-point polarizabilities α_{PQ}. They are symmetric with respect to interchange of P and Q, so there are only $N(N+1)/2$ distinct values.

These point-to-point polarizabilities can then be fitted to a model. A molecular distributed polarizability α_{tu}^{ab} describes the response of electric moment Q_u at site b to an electric field V_t at site a, or vice versa. The labels t and u represent multipole components 00, 10, 11c, 11s, 20, ...; Q_{00} is the charge, Q_{10}, Q_{11c}, and Q_{11s} are the z, x, and y components of the dipole moment, Q_{20} is a component of the quadrupole moment, and so on.[1] V_{00} is the electrostatic potential, $-V_{10}$, $-V_{11c}$, and $-V_{11s}$ are the z, x, and y components of the electric field, $-V_{20}$ is a component of the field gradient, and so on. A distributed polarizability model for a molecule comprises a collection of such polarizabilities. We discuss the form of such models in more detail in Sec. III.

A point charge at P produces fields at a given by functions T_{0t}^{Pa}, where $T_{00}^{Pa} = (4\pi\epsilon_0|P-a|)^{-1}$, and the T_{0t}^{Pa} are derivatives of this quantity with respect to the position a of site a. Detailed formulas are in Appendix F of Ref. 1. Similarly, an induced moment ΔQ_u^b at site b produces a potential $\Delta Q_u^b T_{u0}^{bQ}$ at the point Q. Consequently, the potential at Q induced by a unit point charge at P acting on the model is

$$\tilde{\alpha}_{PQ} = -\sum_{abtu} T_{0t}^{Pa} \alpha_{tu}^{ab} T_{u0}^{bQ}, \qquad (2)$$

where the tilde indicates that this is the response predicted by the model rather than the response α_{PQ} calculated ab initio. The polarizabilities included in the model depend on a number of parameters, which are to be adjusted to give the best agreement, in a least-squares sense, between the $\tilde{\alpha}_{PQ}$ and the α_{PQ}. This is normally a linear optimization problem.

Our method has a significant advantage over the use of induction energies only (i.e., essentially just the α_{PQ} with $P=Q$). The sum in Eq. (2) is a linear combination of the products $T_{0t}^{Pa}(p-a)T_{u0}^{bQ}(q-b)$, where p and q are the positions of points P and Q, and a and b are the positions of sites a and b. Viewed as functions of the polar coordinates of $p-a$ and $q-b$ for fixed a and b, these products form a set of orthogonal functions over the polar angles. This means that if the model comprises a set of polarizabilities referred to a single site, and the data points span one or more spherical shells surrounding that site, the value of any particular polarizability is independent of any other polarizabilities that may be included in the model or omitted from it. In contrast, the expansion functions used to fit the induction energies are $T_{0t}^{Pa}(p-a)T_{u0}^{bP}(p-b)$, which are functions of the single coordinate p for fixed a and b and are not orthogonal.

The orthogonality of our expansion functions is only strictly true for a single-site model, and for a set of points that spans one or more shells of fixed radius. If the points are taken over a region that is not spherical but follows the shape of the molecule, the variation in distance will spoil the orthogonality. Nevertheless, we expect that the overall molecular polarizabilities for any distributed-polarizability model will be reasonably similar. The inclusion or exclusion of higher-rank polarizabilities will not affect the overall dipole–dipole polarizability, but in order to explore distributed dipole–dipole polarizability models we have to include the higher polarizabilities and fit the calculated data as well as possible.

The sum of squares of the residuals, $\tilde{\alpha}_{PQ} - \alpha_{PQ}$, is to be minimized with respect to the polarizability parameters. This leads to a set of n linear equations for n parameters, easily solved if the number of parameters is not too great; because of the approximate orthogonality of the fitting functions, the set of equations remains well-conditioned, and therefore, the problem of needing to invert a near-singular matrix in the least-squares procedure[11] does not arise, and it is not necessary to use methods such as the statistical approach adopted by Dehez et al.[10] in order to avoid numerical instabilities. We have found it sufficient to start with a simple model and to refine it by including further polarizabilities, rejecting them if they do not lead to significant improvement in the sum of squares.

Chipot et al.[23] have discussed the choice of positions for the perturbing point charges, and we have followed their suggestions in this respect. We have used a random selection of points, rejecting any that are within a specified distance of any nucleus (usually two van der Waals radii), or that are too far away from every atom (usually four van der Waals radii).

We have explored the numerical uncertainty in the results by statistical analysis of the data. For each molecule, we carried out several calculations, each calculation involving a different batch of 500 random perturbation points. The

fitted polarizability parameters for any particular model vary from batch to batch, and we have carried out independent calculations on each batch and calculated the standard deviation of the parameter values by standard methods. This provides an estimate of the numerical variability of the tabulated results, which were obtained by combining the data from all batches in a single fit. The estimated numerical errors are typically in the region of 1% or 0.1% of the polarizability values, and are negligible compared with errors due to limitations of basis set or functional.

III. POLARIZABILITY MODELS

In the calculation of distributed polarizabilities *ab initio*, the molecule is partitioned into regions, each with its own origin or reference point.[6] A distributed polarizability α_{tu}^{ab} describes the multipole moment Q_u^b (component u at the origin of region b) induced by the t component of the applied field at the origin of region a, i.e., $-V_t^a$. The partitioning is arbitrary; typically there is a region associated with each atom, with its origin at the nucleus. Every point in space is associated with one of the regions; for example, it may be associated with atom a if it is closer to the nucleus of that atom than to any other. Bader partitioning has also been used.[19,24,25] Partitioning in basis function space has been tried, but it leads to numerical instabilities.[7]

This formulation leads to nonlocal polarizabilities, where a field in one region causes a response in another. In particular, it leads to charge-flow polarizabilities, where a potential difference between two regions leads to a flow of electron charge between them. While these features are entirely reasonable and natural in physical terms, they lead to a very complicated description. The associated dispersion coefficients depend on integrals of the form[26,27]

$$\int \alpha_{tu}^{ab}(i\omega)\alpha_{vw}^{cd}(i\omega)d\omega, \tag{3}$$

and the dispersion energy then involves a sum over these—that is, a fourfold sum over region and a fourfold sum over multipole components. In the case of charge-flow polarizabilities, there are associated dispersion terms proportional to R^{-n} with $2\leq n\leq 5$, and the normal R^{-6} behavior results from a cancellation between these terms. Although conservation of charge guarantees that all terms in R^{-n} with $n\leq 6$ cancel out for large R, there are many such terms and they may be large at typical intermolecular distances, so the cancellation is likely to lead to unacceptable numerical errors.

Consequently, a simpler polarizability model is needed. In earlier work[28] it was shown that the nonlocal description can be transformed to a local description, in which the only polarizabilities α_{tu}^{ab} that occur are those with $a=b$. In the present work, therefore, we are concerned mainly with local polarizability models, though we have investigated nonlocal models to verify that they do not give a significantly better description.

One requirement of a polarizability description is that the induction energy must always be negative for molecules in their ground states. This means that the conventional molecular dipole–dipole polarizability must be positive definite:

when expressed in principal axes the components α_{xx}, α_{yy}, and α_{zz} must all be positive. Similar requirements apply to the higher-rank polarizabilities. The same requirement holds for a distributed-polarizability model; that is, the α_{tu}^{ab}, viewed as a matrix with rows labeled by at and columns by bu, must be positive definite. Small negative eigenvalues may be tolerable in practice, on the grounds that a field leading to positive induction energies could not be realized; it would require all components of the field at all atoms to be zero or small except for the one associated with the negative eigenvalue. Nevertheless, they are clearly better avoided.

IV. DISPERSION COEFFICIENTS

In order to determine dispersion coefficients, it is convenient to use polarizabilities expressed in irreducible tensor form. The general formula is

$$\alpha_{LK(ll')}=\sum_{kk'} \langle ll'kk'|LK\rangle\alpha_{ll'kk'}$$

$$=\sum_{kk'} (-)^{l-l'-K}(2L+1)^{1/2}\begin{pmatrix} l & l' & L \\ k & k' & -K \end{pmatrix}\alpha_{ll'kk'}.$$

For dipole–dipole polarizabilities this leads to the isotropic polarizability

$$\alpha_{00(11)}=-\sqrt{3}\,\bar{\alpha}=-\sqrt{\tfrac{1}{3}}(\alpha_{xx}+\alpha_{yy}+\alpha_{zz}), \tag{4}$$

while the anisotropic part of the polarizability gives

$$\alpha_{20(11)}=\sqrt{\tfrac{2}{3}}\Delta\alpha=\sqrt{\tfrac{2}{3}}[\alpha_{zz}-\tfrac{1}{2}(\alpha_{xx}+\alpha_{yy})], \tag{5a}$$

$$\alpha_{21c(11)}=\sqrt{2}\alpha_{xz}, \tag{5b}$$

$$\alpha_{21s(11)}=\sqrt{2}\alpha_{yz}, \tag{5c}$$

$$\alpha_{22c(11)}=\sqrt{\tfrac{1}{2}}(\alpha_{xx}-\alpha_{yy}), \tag{5d}$$

$$\alpha_{22s(11)}=\sqrt{2}\alpha_{xy}. \tag{5e}$$

Only the components $\alpha_{00(11)}$, $\alpha_{20(11)}$, and $\alpha_{22c(11)}$ may be nonzero for the case of a molecule or site with at least C_{2v} symmetry, and only $\alpha_{00(11)}$ and $\alpha_{20(11)}$ may be nonzero in axial symmetry.

The dispersion coefficients may now be derived from the polarizabilities at imaginary frequency via the integrals of Eq. (3).[1] When a local polarizability model is used, the dispersion energy can be expressed in the form

$$U_{\text{disp}}=-\sum_n \sum_{ab} \sum_{L_aL_bJ} \sum_{K_aK_b} C_n^{ab}(L_aL_bJ,K_aK_b)R^{-n}\bar{S}_{L_aL_bJ}^{K_aK_b}, \tag{6}$$

where a and b label sites in the interacting molecules, and $\bar{S}_{L_aL_bJ}^{K_aK_b}$ is an S function describing the orientational behavior.[1] The index J takes values between $|L_a-L_b|$ and L_a+L_b, with the proviso that L_a+L_b+J is even. The dipole–dipole dispersion coefficients are[29]

$$C_6^{ab}(000,00) = 2W \int_0^\infty du\, \alpha_{00(11)}^a(iu)\alpha_{00(11)}^b(iu), \qquad (7)$$

$$C_6^{ab}(022,0K) = -\sqrt{2}W \int_0^\infty du\, \alpha_{00(11)}^a(iu)\alpha_{2K(11)}^b(iu), \qquad (8)$$

$$C_6^{ab}(202,K0) = -\sqrt{2}W \int_0^\infty du\, \alpha_{2K(11)}^a(iu)\alpha_{00(11)}^b(iu), \qquad (9)$$

$$C_6^{ab}(220,KK') = \frac{1}{5}W \int_0^\infty du\, \alpha_{2K(11)}^a(iu)\alpha_{2K'(11)}^b(iu), \qquad (10)$$

$$C_6^{ab}(222,KK') = \frac{2}{7}W \int_0^\infty du\, \alpha_{2K(11)}^a(iu)\alpha_{2K'(11)}^b(iu), \qquad (11)$$

$$C_6^{ab}(224,KK') = \frac{108}{35}W \int_0^\infty du\, \alpha_{2K(11)}^a(iu)\alpha_{2K'(11)}^b(iu), \qquad (12)$$

where $W = \hbar/(2\pi(4\pi\epsilon_0)^2)$.

Higher dispersion coefficients may be calculated by similar methods. The isotropic part of the quadrupole–quadrupole polarizability is

$$\alpha_{00(22)} = -\sqrt{\frac{1}{5}} \sum_k \alpha_{2k,2k}; \qquad (13)$$

in the sum, k takes the values 0, $1c$, $1s$, $2c$, and $2s$. Dispersion terms arising from the quadrupole–quadrupole polarizability include

$$-R^{-8}C_8^{ab}(000,00)\bar{S}_{000}^{00}, \qquad (14)$$

and

$$-R^{-10}C_{10}^{ab}(000,00)\bar{S}_{000}^{00}, \qquad (15)$$

where the dispersion coefficients are

$$C_8^{ab}(000,00) = -\sqrt{15}W\left\{ \int \alpha_{00(11)}^a(iu)\alpha_{00(22)}^b(iu)du \right.$$
$$\left. + \int \alpha_{00(22)}^a(iu)\alpha_{00(11)}^b(iu) \right\}, \qquad (16)$$

$$C_{10}^{ab}(000,00) = 14W \int \alpha_{00(22)}^a(iu)\alpha_{00(22)}^b(iu)du. \qquad (17)$$

There are many anisotropic contributions to these higher terms also, but our calculations do not at present provide sufficiently well-characterized data to evaluate them reliably.

V. RESULTS AND DISCUSSION

A. Carbon dioxide

1. Polarizabilities

For CO_2 we used a Sadlej basis[30,31] and the B3LYP functional. Calculations were carried out using 24 batches of 500 points each. We should note at the outset that although the Sadlej basis is designed to give good results for properties such as polarizabilities, the higher-rank polarizabilities (dipole–quadrupole and quadrupole–quadrupole) require a greater flexibility in the basis set and are likely to be underestimated. Furthermore, there is no reason to expect the B3LYP functional to give particularly good results for polarizabilities. Our concern here is to demonstrate the general features of our approach.

Table I gives results for a selection of polarizability models. Because of the axial symmetry, the only nonzero components are $\alpha_{lk,l'k'}$ with $k=k'$; moreover, $\alpha_{lkc,l'kc} = \alpha_{lks,l'ks}$. Model A is a single-site description, with dipole polarizabilities only at the center of the molecule. Model B comprises anisotropic dipole polarizabilities on each atom; model C adds charge flow between C and O, and model D adds dipole–quadrupole polarizabilities on O. (The dipole–quadrupole polarizability on C must be zero as it is a centrosymmetric site.) Model E includes quadrupole–quadrupole polarizabilities on each atom, but no dipole–quadrupole terms. This leads to a model that is not positive definite; model F constrains the offending terms in the carbon quadrupole–quadrupole polarizability to be zero.

The first point to notice is that the overall molecular dipole–dipole polarizability is very insensitive to the model. For all but the single-site model, which fits the data rather poorly, the values of the dipole–dipole polarizability obtained by the fitting procedure are very close to the values obtained from a conventional CADPAC polarizability calculation with the same basis set and functional, namely $\bar{\alpha} \approx 16.9$ and $\Delta\alpha \approx 13.0$ a.u. For comparison, the experimental values are $\bar{\alpha} = 17.81$ (Ref. 32) and $\Delta\alpha = 14.62$.[33]

Including higher-rank polarizabilities makes very little difference to the molecular dipole–dipole polarizabilities. This confirms the point made above that by exploring the response at many different points to the perturbing potential from a single point charge, it becomes possible to fit the different polarizabilities independently of each other. The values of the total molecular quadrupole–quadrupole polarizability are meaningless for the simpler models, which do not have enough flexibility to describe it, but the results for models E and F are close to the values obtained from a conventional calculation, namely $\alpha_{2020} = 239.8$, $\alpha_{21c21c} = 213.4$, and $\alpha_{22c22c} = 33.9$ a.u. Model E is slightly better here than model F, where the C atom quadrupole–quadrupole polarizability has been constrained to avoid negative polarizabilities.

The individual atomic dipole–dipole terms are more sensitive to the inclusion of higher-rank terms. Inclusion of charge-flow terms reduces the longitudinal component of the carbon and oxygen polarizabilities, but does not improve the quality of fit, showing that when explicit charge flow is excluded $\alpha_\parallel^C$ and $\alpha_\parallel^O$ pick up its effects satisfactorily.

The inclusion of higher-rank polarizabilities changes the atomic dipole polarizabilities quite significantly. Dipole–quadrupole polarizabilities on oxygen (model D) have little effect on the quality of fit, but quadrupole–quadrupole polarizabilities on C and O improve it substantially (model E). Unfortunately, however, some of the carbon quadrupole polarizabilities are negative for this model, which is physically unsatisfactory. Forcing these terms to be zero gives a more

TABLE I. Some polarizability models for CO_2. The r.m.s. fitting error and the maximum error are expressed as percentages of the range of response values. Values are in atomic units, i.e., $4\pi\epsilon_0 a_0^n$ for the polarizabilities, where $n=3$ for dipole–dipole, 4 for dipole–quadrupole, and 5 for quadrupole–quadrupole; and $E_h a_0^6$ for the C_6 dispersion coefficient.

| | Sites | | | α_{tu}^{ab} | | | | | |
	ab	t	u	A	B	C	D	E	F
	CO	00	00	---	---	−0.22	---	---	---
$\alpha_\parallel^C$	CC	10	10	24.96	12.44	11.03	11.89	12.25	12.45
$\alpha_\perp^C$	CC	11c	11c	12.26	0.56	0.56	2.45	2.22	2.21
	CC	20	20	---	---	---	---	−21.64	---
	CC	21c	21c	---	---	---	---	29.56	29.50
	CC	22c	22c	---	---	---	---	−4.33	---
$\alpha_\parallel^O$	OO	10	10	---	6.56	6.25	6.84	6.66	6.55
$\alpha_\perp^O$	OO	11c	11c	---	6.01	6.01	5.07	5.15	5.15
	OO	10	20	---	---	---	−0.62	---	---
	OO	11c	21c	---	---	---	1.83	---	---
	OO	20	20	---	---	---	---	7.50	0.21
	OO	21c	21c	---	---	---	---	19.52	19.54
	OO	22c	22c	---	---	---	---	18.94	16.89
$\bar{\alpha}$	Total			16.49	16.91	16.91	16.91	16.86	16.86
$\Delta\alpha$	Total			12.70	12.98	12.99	12.98	13.05	13.04
	Total	20	20	---	---	---	---	241.13	244.79
	Total	21c	21c	---	---	---	---	212.26	212.11
	Total	22c	22c	---	---	---	---	33.55	33.71
Fitted parameters				2	4	5	6	10	8
r.m.s. error				0.989	0.171	0.171	0.171	0.034	0.038
Maximum error				24.3	3.37	3.37	3.73	0.82	1.11
Isotropic C_6				141.8	149.4	149.3	149.5	148.6	148.5

sensible model, though the fit is not quite so good.

We see then that almost any model with sufficient flexibility in the atomic dipole polarizabilities is capable of reproducing the overall molecular polarizability, but that if we want an accurate account of the local response to a nonuniform field we need to include higher polarizabilities.

2. Dispersion coefficients

Dispersion coefficients have been calculated by the methods described above, using the frequency-dependent polarizabilities. Table I gives the isotropic C_6 coefficient for each of the models shown. For all but the single-site model, it is close to the value of 148.86 a.u. obtained from a conventional dispersion coefficient calculation using CADPAC with the same basis set and functional. The anisotropic atom–atom coefficients for model F are shown in Table II. The dispersion energy is given by Eq. (6). For a linear molecule, only terms with $K_a=K_b=0$ occur.

TABLE II. C_6 atom–atom dispersion coefficients for CO_2, derived from the frequency-dependent polarizabilities of model F. Values are in atomic units, hartree bohr (Ref. 6).

Atoms: L_aL_bJ,K_aK_b	CC	CO $C_6^{ab}(L_aL_bJ,K_aK_b)$	OO
000, 00	13.08	15.34	18.52
022, 00	6.92	1.14	1.33
202, 00	6.92	7.86	1.33
220, 00	0.76	0.12	0.02
222, 00	1.08	0.17	0.03
224, 00	11.70	1.87	0.31

The oxygen–oxygen dispersion is close to isotropic, since the oxygen polarizability is nearly isotropic. The carbon–carbon dispersion, however, is highly anisotropic. The distance dependence of all of these terms is R^{-6}, but the angular form of the most important anisotropic terms is the same as some of the terms in the electrostatic interaction: the $C_6(022,00)$ and $C_6(202,00)$ terms have an angular dependence like the charge–quadrupole and quadrupole–charge electrostatic interactions, respectively, while the $C_6(224,00)$ term has an angular dependence like the quadrupole–quadrupole electrostatic interaction. Consequently they can be calculated easily by any program, such as ORIENT,[34] which can handle the anisotropic electrostatic interactions. The remaining terms, $C_6(220,00)$ and $C_6(222,00)$, do not have the form of any electrostatic interaction, but they are much smaller and can be neglected without much loss of accuracy. The ORIENT program does provide for them, however, in addition to the electrostatic terms.

VI. WATER

A. Polarizabilities

Table III shows results for water, obtained from 27 batches of 500 points calculated using the Sadlej basis and the B3LYP functional. The O atom and molecular polarizabilities are described in terms of axes with z along the C_2 symmetry axis and y perpendicular to the molecular plane; for the H atoms local axes are used, with z along the O–H bond and y perpendicular to the molecular plane.

The same characteristics emerge as for CO_2. The overall dipole polarizability is almost the same in every case, though the quality of the fit is not as good for the simpler models,

TABLE III. Some polarizability models for H_2O. The r.m.s. fitting error and the maximum error are expressed as a percentage of the range of response values. Values are in atomic units, i.e., $4\pi\epsilon_0 a_0^n$, where $n=3$ for dipole–dipole, 4 for dipole–quadrupole, and 5 for quadrupole–quadrupole.

| | Sites | | | α_{tu}^{ab} | | | | |
	ab	t	u	A	B	C	D	E
α_{zz}^{O}	OH	00	00	---	−1.26	---	---	---
α_{xx}^{O}	OO	10	10	9.91	6.83	7.01	8.99	9.10
α_{yy}^{O}	OO	11c	11c	10.13	5.12	6.74	9.48	8.90
	OO	11s	11s	9.55	9.55	7.31	9.24	9.14
	OO	10	20	---	---	---	−4.03	−4.13
	OO	10	22c	---	---	---	−0.45	−0.68
	OO	11c	21c	---	---	---	−6.05	−5.15
	OO	11s	21s	---	---	---	−3.49	−3.31
	OO	20	20	---	---	---	36.23	35.71
	OO	21c	21c	---	---	---	37.65	35.71
	OO	21s	21s	---	---	---	34.65	35.71
	OO	22c	22c	---	---	---	37.79	35.71
	OO	22s	22s	---	---	---	38.60	35.71
	OO	20	22c	---	---	---	−1.04	---
α_{zz}^{H}	HH	10	10	---	---	2.33	0.99	1.10
α_{xx}^{H}	HH	11c	11c	---	---	0.92	−0.12	---
α_{zx}^{H}	HH	10	11c	---	---	0.02	0.18	---
α_{yy}^{H}	HH	11s	11s	---	---	1.13	0.16	0.20
α_{zz}	Total	10	10	9.91	9.95	9.96	9.94	9.93
α_{xx}		11c	11c	10.13	10.28	10.29	10.27	10.27
α_{yy}		11s	11s	9.55	9.55	9.58	9.55	9.55
Fitted parameters				3	4	7	17	10
r.m.s. error				0.674	0.424	0.247	0.086	0.090
Maximum error				19.0	12.8	6.4	2.8	3.1
Isotropic C_6				46.20	46.76	46.90	46.65	46.66

especially model A, which simply has a dipole–dipole polarizability at oxygen. Adding a charge-flow term (model B) improves the fit somewhat, but a greater improvement is achieved with dipole polarizabilities at the H atoms as well as O (model C). Adding dipole–quadrupole and quadrupole–quadrupole polarizabilities on oxygen (model D) improves the fit significantly, but the H-atom polarizabilities are no longer positive definite. Suppressing the hydrogen α_{xx} and α_{xz} gives a positive definite description, and the quadrupole–quadrupole polarizability on oxygen can be made isotropic to reduce the number of fitted parameters. These changes (model E) make the fit only slightly worse.

The constancy of the overall polarizability values may be compared with the variation seen in the paper by Celebi et al.,[11] where the distributed polarizabilities were fitted to induction energy values. Leaving aside their model A, which includes charge-flow polarizabilities only and is not expected to give a sensible result, their models B and C give reasonably consistent values for the dipole polarizability, though not as constant as ours; but, their addition of the quadrupole–quadrupole polarizability on O changes the dipole–dipole values by as much as 18%. Moreover, it is only when the quadrupole–quadrupole polarizability is included, in their model E, that their dipole–dipole polarizabilities approach the values obtained in a conventional calculation.

In contrast, our dipole–dipole polarizabilities are almost constant, and even for the simplest model are close to the values obtained from a conventional polarizability calculation using the same basis set and functional, which are α_{zz} = 9.96, α_{xx} = 10.31, and α_{yy} = 9.61.

An interesting feature emerges if we take a single-site model with the site not at the O atom but displaced along the symmetry axis towards the hydrogen atoms. The dipole polarizability and the C_6 coefficient are formally independent of origin, and indeed, they change only slightly. However, the r.m.s. error is reduced from 0.67% to 0.57% for an origin at the center of mass, 0.12 bohr from the O atom, and to 0.44% for an origin 0.35 bohr from the O atom.

B. Dispersion coefficients

Table IV gives atom–atom dispersion coefficients for water, derived from the frequency-dependent polarizabilities of model E. The low symmetry of the atom sites, C_{2v} for O and C_s for H, leads potentially to a large number of nonzero dispersion coefficients. However, the choice of local axes for H with z along the O–H bond makes the xz component of the polarizability very small, and it is suppressed altogether in model E. Moreover, the anisotropy of the molecular polarizability is quite small, and indeed the local O atom polarizabilities are nearly isotropic, while the H-atom polarizabilities are small.

Our analysis justifies the common assumption that a single-site model with isotropic polarizability and dispersion is adequate for all but the most accurate work, but it suggests that both should be attached to a site on the symmetry axis about 0.35 bohr from the O atom rather than at the O atom or the center of mass.

TABLE IV. C_6 atom–atom dispersion coefficients for H_2O, derived from the frequency-dependent polarizabilities of model E. Values are in atomic units, hartree bohr (Ref. 6).

Atoms: L_aL_bJ,K_aK_b	OO	OH $C_6^{ab}(L_aL_bJ,K_aK_b)$	HH
000, 00	39.04	1.82	0.08
202, 00	0.14	0.01	0.07
202, 2c0	0.51	0.02	−0.01
022, 00	0.14	1.42	0.07
022, 02c	0.51	−0.23	−0.01
220, 00	0.00	0.00	0.01
222, 00	0.00	0.00	0.01
224, 00	0.00	0.01	0.16
220, 02c	0.00	0.00	0.00
222, 02c	0.00	0.00	0.00
224, 02c	0.01	0.00	0.02
220, 2c0	0.00	0.00	0.00
222, 2c0	0.00	0.01	0.00
224, 2c0	0.01	0.05	0.02
220, 2c2c	0.00	0.00	0.00
222, 2c2c	0.00	0.00	0.00
224, 2c2c	0.04	0.01	0.01

VII. BENZENE

A. Polarizabilities

For benzene we used the Sadlej basis and the B3LYP functional, and calculated responses for 20 batches of 500 points. The response values were fitted to a number of polar-izability models, and some results are shown in Table V. We also carried out calculations using the PBE0 functional; the results were similar but the overall polarizabilities somewhat smaller in magnitude.

A single-site model (model A), with dipole polarizabilities at the center of mass, gives a poor fit. Adding quadrupole polarizabilities at the center of mass does not improve matters very much. Distributed-polarizability models are better; with dipole polarizabilities at the carbon atoms only (model B), the polarizability values are already in good agreement with the experimental values, $\bar{\alpha}=70.3$ and $\Delta\alpha=-38$, and are close to the values, $\bar{\alpha}=70.1$ and $\Delta\alpha=-37.8$, obtained from a conventional *ab initio* polarizability calculation using the same basis set and functional. Adding sites on the H atoms (model C) improves the fit significantly. Adding quadrupole polarizabilities on the carbon atoms (model D) gives a further substantial improvement, but the atomic polarizabilities now are not positive definite. The components listed are all those that are allowed by the local C_{2v} symmetry of the carbon site. Closer investigation shows that it is the off-diagonal α_{2022c}^{C} term that is responsible for the lack of positive definiteness; it can be forced to zero with only a slight reduction in the quality of the fit.

It is evident that there is considerable variation between models in the values of the atomic dipole–dipole polarizability components, but that the overall dipole–dipole polarizabilities and isotropic dispersion coefficients hardly vary at all. One should not expect a definitive model for the *distrib-*

TABLE V. Some polarizability models for benzene. The r.m.s. fitting error and the maximum error are expressed as a percentage of the range of response values. Values are in atomic units, i.e., $4\pi\epsilon_0 a_0^n$, where $n=3$ for dipole–dipole, 4 for dipole–quadrupole, and 5 for quadrupole–quadrupole. The local axes for the C and H atoms have z perpendicular to the molecular plane and y along the C–H bond. The X site for model A is the center of mass.

	Sites			α_{tu}^{ab}				
	ab	t	u	A	B	C	D	E
$\alpha_{\parallel}^{X}$	XX	10	10	42.21	---	---	---	---
$\alpha_{\perp}^{X}$	XX	11c	11c	80.72	---	---	---	---
α_{zz}^{C}	CC	10	10	---	7.45	4.82	6.65	6.29
α_{xx}^{C}	CC	11c	11c	---	9.53	19.50	5.80	6.29
α_{yy}^{C}	CC	11s	11s	---	18.13	4.23	16.55	16.35
	CC	10	21s	---	---	---	2.92	1.31
	CC	11c	22s	---	---	---	−21.06	−8.06
	CC	11s	22c	---	---	---	−3.21	11.45
	CC	11s	20	---	---	---	11.52	15.27
	CC	20	20	---	---	---	26.69	22.24
	CC	20	22c	---	---	---	40.54	---
	CC	21c	21c	---	---	---	78.34	88.60
	CC	21s	21s	---	---	---	89.04	77.69
	CC	22c	22c	---	---	---	43.17	52.94
	CC	22s	22s	---	---	---	79.35	96.34
α_{zz}^{H}	HH	10	10	---	---	2.78	0.81	1.16
α_{xx}^{H}	HH	11c	11c	---	---	−0.09	2.63	1.16
α_{yy}^{H}	HH	11s	11s	---	---	3.90	2.49	3.67
$\bar{\alpha}$	Total			67.89	70.22	70.29	69.86	69.85
$\Delta\alpha$	Total			−38.52	−38.26	−37.08	−37.71	−37.70
	Total	20	20	---	---	---	816.8	818.5
	Total	21c	21c	---	---	---	1159.3	1159.4
	Total	22c	22c	---	---	---	2294.7	2290.4
Fitted parameters				2	3	6	16	13
r.m.s. error				1.436	0.490	0.258	0.061	0.065
Maximum error				35.0	13.3	10.3	3.5	5.0
Isotropic C_6				1658.3	1788.9	1790.0	1768.1	1767.9

TABLE VI. C_6 atom–atom dispersion coefficients for benzene, derived from the frequency-dependent polarizabilities of model E, and referred to local axes with z along the C–H bond. Values are in atomic units, hartree bohr (Ref. 6).

L_aL_bJ, K_aK_b	$C_6^{ab}(L_aL_bJ, K_aK_b)$		
Atoms:	CC	CH	HH
000, 00	31.42	7.84	2.00
202, 00	9.30	2.26	0.56
022, 00	9.30	2.31	0.56
220, 00	0.57	0.14	0.04
222, 00	0.81	0.20	0.05
224, 00	8.79	2.20	0.55

uted polarizabilities; the partition between atoms of the polarization response is largely arbitrary. (Bader's atoms in molecules approach[35] would give a less arbitrary partitioning, but so far that approach appears to have been used only to dissect the response of the molecule to a uniform field, either static[24] or dynamic.[19]) The arbitrariness provides an opportunity to simplify the model without losing accuracy. For example, the dipole polarizability components perpendicular to the C–H bond in model D can be constrained to be equal, giving a fit that is almost as good, but with two fewer fitted parameters. (Three fewer if we also constrain α_{2022c}^C to be zero; this is model E.) The advantage of this apparently arbitrary procedure is that it leads to a much simpler dispersion model, as we show below.

The total molecular quadrupole–quadrupole polarizability is shown in Table V for those models that have enough flexibility to describe it adequately. The values of $\alpha_{20,20}$, $\alpha_{21c,21c}$, and $\alpha_{22c,22c}$ are close to the values obtained from a conventional CADPAC calculation: $\alpha_{20,20} = 822.0$, $\alpha_{21c,21c} = 1157.7$, and $\alpha_{22c,22c} = 2296.0$.

B. Dispersion coefficients

Table V shows the isotropic molecular C_6 dispersion coefficient for each of the benzene models considered. We can see that, like the total molecular polarizability, it is well described by all but the simplest of the models. The distributed description of dispersion, like the distributed model of the polarizability, varies a good deal between models. In this case we can take advantage of the flexibility of the distributed model to base the dispersion coefficients on model E, where the polarizability of both carbon and hydrogen has axial symmetry about the C–H bond. We redefine the local z axis to lie along the C–H bond, and obtain the dispersion coefficients listed in Table VI. Because of the effective axial symmetry, the same components are nonzero as for CO_2, and the same comments apply. A dispersion model based on model D has more than twice as many different nonzero C_6 coefficients, and would be more complicated to use while offering very little advantage.

VIII. ETHANE

For ethane (Table VII) we used 25 batches of 500 points, calculated using the Sadlej basis and the B3LYP functional. As for the other cases described above, the overall molecular dipole polarizability is in good agreement with the results of a conventional CADPAC calculation ($\bar{\alpha} = 29.0$ a.u., $\Delta\alpha = 3.84$ a.u.), and in reasonable agreement with the experimental data: $\bar{\alpha} = 30$ a.u.,[36] $\Delta\alpha = 4.5$ a.u.[37]

However, model A, with dipole polarizabilities on the C atoms only, gives a rather poor fit to the response data. Model B, with C atom polarizabilities up to quadrupole, is significantly better, but a greater improvement is achieved with dipole polarizabilities only on both C and H sites (model C). The best result is obtained with dipole polariz-

TABLE VII. Some polarizability models for ethane. The r.m.s. fitting error and maximum error are expressed as percentages of the range of response values. Values are in atomic units, i.e., $4\pi\epsilon_0 a_0^n$, where $n = 3$ for dipole–dipole, 4 for dipole–quadrupole, and 5 for quadrupole–quadrupole. For the C atoms, the local z axis is directed outwards along the CC bond and the y axis parallel to one of the molecular C_2 dihedral symmetry axes. For the H atoms, the local z axis is along the bond and the y axis is perpendicular to the HCC plane.

	Sites			α_{tu}^{ab}				
	ab	t	u	A	B	C	D	E
$\alpha_\parallel^C$	CC	10	10	15.79	15.68	6.40	9.35	10.11
$\alpha_\perp^C$	CC	11c	11c	13.70	13.72	2.97	5.03	5.22
	CC	10	20	---	−5.09	---	0.78	−0.13
	CC	11c	21c	---	9.78	---	1.49	3.68
	CC	11c	22c	---	−16.92	---	−6.73	3.65
	CC	20	20	---	95.70	---	11.25	8.44
	CC	21c	21c	---	132.64	---	46.13	46.13
	CC	22c	22c	---	122.32	---	22.85	25.00
	CC	21c	22c	---	2.34	---	10.19	8.62
α_{zz}^H	HH	10	10	---	---	4.87	3.78	4.28
α_{xx}^H	HH	11c	11c	---	---	2.58	1.71	1.62
α_{yy}^H	HH	11s	11s	---	---	2.82	2.43	1.62
α_{xz}^H	HH	11c	10	---	---	0.55	0.38	---
$\bar{\alpha}$	Total			28.79	28.75	28.77	28.77	28.77
$\Delta\alpha$				4.17	3.92	3.88	3.91	3.91
Fitted parameters				2	9	6	13	11
r.m.s. error				0.637	0.159	0.087	0.051	0.057
Maximum error				24.1	9.0	2.4	1.4	2.1
Isotropic C_6				361.1	359.3	356.0	359.8	359.8

TABLE VIII. C_6 atom–atom dispersion coefficients for ethane, derived from the frequency-dependent polarizabilities of model E, and referred to local axes with z along the C–C bond for the C site and along the C–H bond for H. Values are in atomic units, hartree bohr (Ref. 6).

Atoms: $L_a L_b J, K_a K_b$	CC	CH $C_6^{ab}(L_a L_b J, K_a K_b)$	HH
000, 00	20.75	7.48	2.70
202, 00	3.61	1.31	0.72
022, 00	3.61	1.97	0.72
220, 00	0.14	0.08	0.04
222, 00	0.20	0.11	0.06
224, 00	2.17	1.18	0.66

abilities on H and polarizabilities up to quadrupole on C (model D). However, as in the case of benzene, a model in which the H-atom polarizabilities are forced to have axial symmetry about the C–H bond (model E) fits almost as well and gives a simpler description of the dipole polarizability and the C_6 dispersion coefficients.

The overall molecular C_6 coefficient is given in Table VII for each model listed, and the anisotropic coefficients are given in Table VIII. Although the molecular polarizability is not very anisotropic, the distributed polarizabilities are more anisotropic, and this is reflected in the C_6 coefficients. As for CO_2, however, the only important terms are $C_6(202,00)$ and $C_6(022,00)$, which have angular behavior like the charge–quadrupole electrostatic interaction, and $C_6(224,00)$, which is like the quadrupole–quadrupole, though all of them have R^{-6} radial dependence. As in the case of benzene, forcing the H-atom polarizabilities to have axial symmetry greatly simplifies the dispersion model.

IX. CONCLUSIONS

We have implemented a modification of the method of Dehez et al.[10] for determining distributed polarizabilities. We calculate the response (the change in electrostatic potential) at each of an array of points around the molecule to a unit point charge at each point in turn, and fit the resulting data to a distributed polarizability model. This modified approach has several advantages. The fitted functions are orthogonal with respect to integration over the angular position of either the perturbing charge or the response point, and although the array of points used follows the shape of the molecule rather than a spherical shell, the underlying orthogonality makes the fitting equations well-conditioned and stable. The molecular dipole polarizability is almost independent of the model used, as long as it is sufficiently flexible, and it agrees very closely with the polarizability obtained from a conventional calculation. However, the fitting procedure gives information about the distribution of polarizability among the atoms of the molecule, and about the anisotropy of the atomic polarizabilities.

Furthermore, the procedure can be applied just as easily to polarizabilities calculated at imaginary frequency, and this provides a route to anisotropic atom–atom dispersion coefficients. In this way we can study the distribution of the mo-

lecular dispersion interaction between the atoms in a polyatomic molecule, and obtain anisotropic atom–atom dispersion coefficients.

In the present work we have concentrated on dipole–dipole polarizabilities and C_6 dispersion coefficients. Higher-rank polarizabilities and dispersion coefficients will require larger basis sets. However, the force fields that are commonly used in molecular simulations generally use only isotropic R^{-6} dispersion terms, and ab initio information about the distribution of dispersion interactions between atoms will be a valuable source of information for refining such force fields.

[1] A. J. Stone, The Theory of Intermolecular Forces, International Series of Monographs in Chemistry (Clarendon, Oxford, 1996).
[2] A. Kumar and W. J. Meath, Mol. Phys. **75**, 311 (1992).
[3] A. Kumar and W. J. Meath, Chem. Phys. **189**, 467 (1994).
[4] A. Kumar, W. J. Meath, P. Bundgen, and A. J. Thakkar, J. Chem. Phys. **105**, 4927 (1996).
[5] S. A. C. McDowell and W. J. Meath, Mol. Phys. **90**, 713 (1997).
[6] A. J. Stone, Mol. Phys. **56**, 1065 (1985).
[7] C. R. Le Sueur and A. J. Stone, Mol. Phys. **78**, 1267 (1993).
[8] A. J. Stone and C. S. Tong, Chem. Phys. **137**, 121 (1989).
[9] S. Nakagawa and N. Kosugi, Chem. Phys. Lett. **210**, 180 (1993).
[10] F. Dehez, C. Chipot, C. Millot, and J. Ángyán, Chem. Phys. Lett. **338**, 180 (2001).
[11] N. Celebi, J. G. Ángyán, F. Dehez, C. Millot, and C. Chipot, J. Chem. Phys. **112**, 2709 (2000).
[12] R. D. Amos, Tech. Rep., University of Cambridge (1995), a suite of quantum chemistry programs developed by R. D. Amos with contributions from I. L. Alberts, J. S. Andrews, S. M. Colwell et al.
[13] R. J. W. Le Fèvre, Adv. Phys. Org. Chem. **3**, 1 (1965).
[14] J. Applequist, J. R. Carl, and K.-K. Fung, J. Am. Chem. Soc. **94**, 2952 (1972).
[15] J. Applequist, Acc. Chem. Res. **10**, 79 (1977).
[16] J. Applequist, J. Chem. Phys. **83**, 809 (1985).
[17] M. B. Ferraro, M. C. Caputo, and P. Lazzeretti, J. Chem. Phys. **109**, 2987 (1998).
[18] J. R. Maple and C. S. Ewig, J. Chem. Phys. **115**, 4981 (2001).
[19] C. Hättig, G. Jansen, B. A. Hess, and J. G. Ángyán, Can. J. Chem. **74**, 976 (1996).
[20] S. M. Colwell, N. C. Handy, and A. M. Lee, Phys. Rev. A **53**, 1316 (1996).
[21] P. J. Stephens, F. Devline, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
[22] C. Van Caillie and R. D. Amos, Chem. Phys. Lett. **328**, 446 (2000).
[23] C. Chipot, F. Dehez, J. Ángyán, C. Millot, M. Orozco, and F. J. Luque, J. Phys. Chem. A **105**, 11505 (2001).
[24] R. F. W. Bader, T. A. Keith, K. M. Gough, and K. E. Laidig, Mol. Phys. **75**, 1167 (1992).
[25] J. G. Ángyán, G. Jansen, M. Loos, C. Hättig, and B. A. Hess, Chem. Phys. Lett. **219**, 267 (1994).
[26] H. B. G. Casimir and D. Polder, Phys. Rev. **73**, 360 (1948).
[27] A. D. McLachlan, Proc. R. Soc. London, Ser. A **271**, 387 (1963).
[28] C. R. Le Sueur and A. J. Stone, Mol. Phys. **83**, 293 (1994).
[29] A. J. Stone and R. J. A. Tough, Chem. Phys. Lett. **110**, 123 (1984).
[30] A. J. Sadlej, Collect. Czech. Chem. Commun. **53**, 1995 (1988).
[31] A. J. Sadlej, Theor. Chim. Acta **79**, 123 (1991).
[32] J. M. St. Arnaud and T. K. Bose, J. Chem. Phys. **71**, 4951 (1979).
[33] I. R. Gentle, D. R. Laver, and G. L. D. Ritchie, J. Phys. Chem. **83**, 3035 (1989).
[34] A. J. Stone, A. Dullweber, O. Engkvist, E. Fraschini, M. P. Hodges, A. W. Meredith, P. L. A. Popelier, and D. J. Wales, ORIENT: A program for studying interactions between molecules, version 4.4, University of Cambridge (2001), enquiries to A. J. Stone, ajs1@cam.ac.uk, URL http://www-stone.ch.cam.ac.uk/.
[35] R. Bader, Atoms in Molecules (Clarendon, Oxford, 1990).
[36] M. P. Bogaard and B. J. Orr, in Molecular Structure and Properties, edited by A. D. Buckingham, International Review of Science, Physical Chemistry, Series Two (Butterworth, London, 1975), Vol. 2, Chap. 5.
[37] F. Baas and K. D. van den Hout, Physica A **95**, 597 (1979).