The Journal of Physical Chemistry A

© Copyright 1998 by the American Chemical Society VOLUME 102, NUMBER 14, APRIL 2, 1998

ARTICLES

Molecular and Atomic Polarizabilities: Thole's Model Revisited

Piet Th. van Duijnen* and Marcel Swart

Department of Chemistry (OMAC), University of Groningen, Nijenborgh 4,
9747 AG Groningen, The Netherlands

Received: November 14, 1997; In Final Form: January 19, 1998

Thole's modified dipole interaction model for constructing molecular polarizabilities from effective, isotropic atomic polarizabilities is reviewed and extended. We report effective atomic polarizabilities for H, C, N, O, S, and the halogen atoms, independent of their chemical environment. They are obtained by fitting the model both to experimental and calculated molecular polarizabilities, the latter to enable one to model ab initio polarizabilities for various basis sets.

Introduction

Atomic and molecular dipole polarizabilities are becoming of increasing importance, e.g., as parameters of "polarizable" force fields, as they are developed presently in the field of mixed quantum/classical—or QM/MM—calculations on extended systems.[1−5] In general, in microscopic descriptions of matter, polarizabilities should be used instead of "effective dielectric constants", because in an arbitrary microscopic collection of point charges and polarizabilities the energy and the forces depend on the nature of the charges, and on the spatial arrangement of the charges and polarizabilities. As a consequence, interactions between like and unlike charges do in general not show the uniform scaling implied in using a (macroscopic!) dielectric "constant".[6,7]

In particular, in classical force fields depending on atomic pair contributions it is common to model the inductive and dispersion interactions by means of isotropic atomic polarizabilities, which in general depend on the "chemical environment". Usually they are then considered as adjustable parameters in fitting procedures, in which they lose their precise physical meaning. For example, in one of the first polarizable force fields, CHARMM,[8] one finds for an aliphatic carbon atom a polarizability of 0.98 Å^3, and for an aliphatic hydrogen 0.1 Å^3. In a simple-minded way, i.e., summing the atomic contributions, this would give for CH_4 a polarizability of 1.4 Å^3 while the experimental value is 2.65 Å^3. Similarly, for CH_3-OH, using the same polarizability for C, and defining O and H as "hydroxyl", the model polarizability would be 2.16 Å^3 versus an experimental value of 3.33 Å^3. Likewise, in recent work of Thompson,[3] the polarizabilities of H and O (0.170 and 0.528 Å^3, respectively) lead to a polarizability for water of 0.70 Å^3 instead of the experimental value, ∼1.5 Å^3.

In the force field for QM/MM or classical calculations, developed in our laboratory,[4,5] we consider molecular polarizabilities not as adjustable parameters but rather as well-defined physical properties of the (sub)system(s) of interest. In these cases, we want polarizabilities as close as possible to their (experimental or calculated) vacuum values. Thole devised a model to calculate molecular polarizabilities[9] from (model) isotropic atomic polarizabilities, although formally no relation exists between these entities. The atomic polarizabilities, without need to define atomic "types" or their "environment", were adjustable parameters, obtained by fitting

---

* Corresponding author. E-mail address: ptvd@chem.rug.nl. Telephone: (+31)(0)50-3634373. Fax: (+31)(0)50-3634296

S1089-5639(98)00221-7 CCC: $15.00 © 1998 American Chemical Society

TABLE 1: Experimental and Fitted Molecular Polarizabilities (au) from Fit to 16 Molecules[a]

molecule	exp	linear model	dev, %	expon model	dev, %	Thole 1	dev, %	Thole 2	dev, %
				Learning Set					
acetone	43.12	42.57	1.27	42.47	1.51	42.44	1.57	42.65	1.10
acetonitrile	30.23	28.80	4.74	29.22	3.34	28.53	5.63	28.61	5.38
carbon monoxide	13.16	13.08	0.64	13.06	0.77	12.28	6.71	12.28	6.67
cyclohexane	74.23	73.45	1.05	72.41	2.46	73.96	0.37	73.89	0.45
cyclopentane	61.75	61.38	0.61	60.81	1.52	61.45	0.49	61.55	0.33
dimethyl ether	35.36	35.27	0.26	35.17	0.53	35.41	0.15	35.36	0.00
p-dioxane	58.04	59.27	2.13	58.57	0.90	59.68	2.82	59.66	2.79
ethane	30.23	29.90	1.09	30.12	0.37	29.79	1.46	30.10	0.45
ethanol	34.28	35.22	2.74	35.07	2.29	35.22	2.75	35.50	3.54
formaldehyde	16.53	17.46	5.66	18.00	8.88	16.91	2.30	17.17	3.67
hydrogen	5.33	5.29	0.74	5.29	0.85	5.06	5.08	5.13	3.80
methanol	22.40	22.58	0.80	22.76	1.62	22.37	0.11	22.54	0.60
nitrogen	11.88	11.98	0.83	11.95	0.62	11.57	2.63	11.61	2.27
oxygen	10.80	10.54	2.43	10.48	3.00	10.02	7.23	10.06	6.88
propane	43.05	42.30	1.74	42.18	2.02	42.40	1.52	42.45	1.41
tert-butyl cyanide	64.72	65.83	1.72	65.15	0.67	66.07	2.09	66.34	2.50
deviation, %	16 mol		1.78 ± 1.47		1.96 ± 1.99		2.68 ± 2.22		2.62 ± 2.17
				Control Set					
acetamide	38.26	38.51	0.67	38.43	0.43	38.27	0.02	38.53	0.71
ethylene oxide	29.89	29.30	1.97	30.11	0.73	29.63	0.89	29.63	0.90
formamide	27.53	25.86	6.05	26.18	4.91	25.43	7.63	25.58	7.11
methane	17.68	17.46	1.26	17.88	1.10	17.12	3.20	17.21	2.67
propanol	47.04	47.40	1.91	47.27	0.49	48.16	2.38	48.39	2.87
water	10.06	10.06	0.00	10.21	1.50	9.61	4.45	9.72	3.36
deviation, %	6 mol		1.98 ± 1.95		1.53 ± 1.56		3.10 ± 2.49		2.94 ± 2.11

[a] Thole 1: Thole's atomic polarizabilities in present geometries. Thole 2: Taken from Thole's article.

TABLE 2: Atomic Polarizabilties (au) Fitted to the Original Learning Set

	present		Thole	
model atom	linear	expon	linear	expon
H	3.5308	3.0588	3.4686	2.8815
C	10.0533	8.7939	9.4614	8.6716
N	7.6430	6.6704	7.4557	6.5256
O	6.2955	5.6480	5.8171	5.3042
a	*1.7823*	*1.9088*	*1.662*	*2.089*

TABLE 3: CPHF Polarizabilities (au) of the (Six) Control Set in Various Basis Sets

molecule	exp	STO	N31	DZV	DZP	TZV	TZP
acetamide	38.26	5.12	25.90	27.46	28.95	30.26	31.08
ethylene oxide	29.89	11.81	21.24	21.93	22.85	23.56	24.17
formamide	27.53	9.98	16.66	17.64	18.82	19.96	20.59
methane	17.68	5.54	11.59	11.96	12.60	12.79	13.40
propanol	47.04	18.43	33.61	34.56	36.24	37.46	38.60
water	10.06	2.39	3.90	4.20	5.09	5.01	5.56

the model to experimental polarizabilities of a small "learning set" of molecules, and proved in many unrelated cases to give molecular polarizabilities in most cases with experimental accuracy.

In the repeated process of migrating to other computer platforms, Thole's original programs and data got lost. Furthermore, we were in need of an extended "learning set" of both experimental and computed polarizabilities, the latter to compare fully quantum chemical with QM/MM and fully MM calculations. To this end we recoded Thole's method, and we reparameterized it. In the following sections we review briefly the method and present results of its application on the polarizabilities of up to 70 different molecules chosen from varying classes.

Thole's Modified Dipole Interactions[9]

For a collection of N point dipole polarizabilities placed in a homogeneous electric field $\mathbf{F}$, the induced dipole moment at point p is given by:

$$\mu_p = \boldsymbol{\alpha}_p[\mathbf{F}_p - \sum_{q\neq p}^{N} \mathrm{T}_{pq}\mu_q] \tag{1}$$

with $\boldsymbol{\alpha}_p$ the atomic polarizability tensor of p and T_{pq} the dipole field tensor:

$$\mathrm{T}_{pq} = \frac{\mathbf{1}}{r_{pq}^3} - \frac{3}{r_{pq}^5}\begin{bmatrix} x^2 & xy & \mathrm{xz} \\ \mathrm{yx} & y^2 & yz \\ \mathrm{zx} & \mathrm{zy} & z^2 \end{bmatrix} \tag{2}$$

where $\mathbf{1}$ is the 3 × 3 unit tensor, r_{pq} is the distance between atoms p and q, and x, y, and z are the Cartesian components of the vector connecting the two.

Equation 1 can be rearranged into the single matrix equation $\mathrm{R}\mathbf{M} = \mathbf{F}$, where $\mathbf{M}$ is a $3N$ vector containing the induced moments, and R, the relay matrix, is a $3N \times 3N$ matrix:

$$\mathrm{R} = \begin{bmatrix} \boldsymbol{\alpha}_1^{-1} & 0 & 0 & 0 \\ 0 & \boldsymbol{\alpha}_2^{-1} & \cdots & \cdots \\ \vdots & \cdots & \ddots & \cdots \\ 0 & 0 & 0 & \boldsymbol{\alpha}_N^{-1} \end{bmatrix} + \begin{bmatrix} 0 & \mathrm{T}_{12} & \cdots & \mathrm{T}_{1N} \\ \mathrm{T}_{21} & 0 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \cdots \\ \mathrm{T}_{N1} & \cdots & \cdots & 0 \end{bmatrix} \tag{3}$$

containing the inverse of the (isotropic) atom polarizability tensors in 3 × 3 diagonal blocks which are coupled by the dipole field tensors T_{pq}.

Inversion of R yields:

$$\mathrm{R}^{-1} = [\boldsymbol{\alpha}^{-1} + \mathrm{T}]^{-1} \equiv \mathrm{A} \tag{4}$$

A maps the linear response of the system to a (uniformly applied) electric field. It is just a polarizability, as is obvious from $\mathbf{M} = \mathrm{A}\mathbf{F}$, albeit in a $3N \times 3N$ "atomic" representation. A can be reduced to a "normal" (3 × 3) molecular polarizability tensor:

TABLE 4: Atomic Polarizabilities from Fit to 16 ab Initio Molecular (CPHF) Polarizabilities

basis set	STO		N31		DZV		DZP		TZV		TZP	
model	linear	expon	linear	expon	linear	expon	linear	expon	linear	expon	linear	expon
H	0.5278	0.5410	1.2388	1.1178	1.2453	1.0989	.5388	1.3333	.2931	1.1550	1.6106	1.3899
C	3.7757	3.8570	7.2605	6.8721	8.2030	7.2352	8.0484	7.2545	9.0813	8.0552	9.0015	7.8995
N	2.1716	2.1945	4.6255	3.9125	4.9165	4.1303	5.4064	4.5343	5.3701	4.4997	5.6740	4.7383
O	1.3761	1.4326	2.7017	2.7382	2.9236	2.8681	3.5749	3.3791	3.2613	3.1986	3.6476	3.4939
a	*1.9331*	*2.4055*	*1.7020*	*2.4966*	*1.7213*	*2.4864*	*1.6898*	*2.4450*	*1.6920*	*2.4297*	*1.6766*	*2.4290*
dev	3.42	5.61	3.61	4.05	4.77	3.80	3.82	3.41	3.67	3.21	3.37	2.85

TABLE 5: Average Percentage of Experimental Polarizability

basis set	STO	N31	DZV	DZP	TZV	TZP
SOP	32.2	55.9	56.1	62.8	62.4	67.0
SOP (Unsöld)	65.3	79.3	82.5	70.5	84.4	74.2
CPHF	35.8	66.2	68.7	73.0	74.7	77.8

TABLE 6: Empirical Atomic Polarizabilities for Some Halogen Atoms

atom	linear	expon
F	3.1685	3.0228
Cl	16.5421	16.8649
Br[a]	22.8464	23.2597
a	*1.7823*	*1.9088*

[a] Fitted to one experimental value.

$$\alpha^{G}_{mn} = \sum_{i,j=1}^{N} (A_{ij})_{mn} \qquad m,n \in \{x,y,z\} \tag{5}$$

or, for that matter, to any partition of N in subgroups by summing corresponding 3 × 3 blocks.

A well-known property of this (so far nonmodified) point dipole model is that it may lead to infinite polarization by the cooperative (head to tail) interaction between two induced dipoles in the direction of the line connecting the two. From the above, one may derive the resulting polarizability components:

$$\alpha_{||} = \frac{\alpha_p + \alpha_q + 4\alpha_p\alpha_q/r^3}{1 - 4\alpha_p\alpha_q/r^6}$$

$$\alpha_{\perp} = \frac{\alpha_p + \alpha_q - 2\alpha_p\alpha_q/r^3}{1 - \alpha_p\alpha_q/r^6} \tag{6}$$

When the distance r between the two points approaches $4(\alpha_p\alpha_p)^{1/6}$, $\alpha_{||}$ goes to infinity and will be negative for shorter distances.

Thole "repaired" this deficiency by modifying the dipole field tensor:

$$(T_{pq})_{ij} = \delta_{ij}r^{-3} - 3x_ix_jr^{-5}$$

$$= (\alpha_p\alpha_q)^{-1/2}(\delta_{ij}u^{-3} - 3u_iu_ju^{-5})$$

$$= (\alpha_p\alpha_q)^{-1/2}\, t_{ij}(\mathbf{u}) \tag{7}$$

where $\mathbf{u} = \mathbf{x}/(\alpha_p\alpha_q)^{1/6}$ and δ_{ij} is the Kronecker delta.

TABLE 7: Effective Atomic Polarizabilities (au) from Fit to 52 Molecules

	empirical		ab initio (mean)		ab initio components	
atom	linear	expon	linear	expon	linear	expon
H	3.5020	2.7927	1.5350	1.3368	0.5974	1.3849
C	10.1756	8.6959	8.7622	7.9421	7.7764	6.5780
N	7.6048	6.5565	5.2663	4.5428	4.7575	2.9855
O	6.3940	5.7494	2.6175	2.5694	2.6442	2.3184
F	2.9413	3.0013	0.5819	1.1109	0.8149	1.4906
S	19.7422	16.6984	15.0931	13.9500	11.0208	10.8651
Cl	16.1145	16.1979	7.0741	9.0475	7.5682	6.7129
Br	22.6671	23.5714	12.1193	15.7347	12.7339	12.3384
I	34.2434	36.9880	19.6286	23.4488	20.2867	20.0081
a	*1.7278*	*2.1304*	*1.6767*	*2.4380*	*1.6623*	*2.8990*
dev, %	3.64	3.76	6.10	5.90	7.62	9.98

In eq 6 we have:

$$T_{pq} = (\alpha_p\alpha_q)^{1/2}t(\mathbf{x}/(\alpha_p\alpha_q)^{1/6}) \tag{8}$$

with t a *shape function*, which does not depend on p and q, but is related to a well-behaved (model) charge distribution $\rho(u)$. Thole originally investigated various forms for the charge distribution of which two survived:

$$\rho(u) = \begin{cases} \dfrac{3}{\pi}\dfrac{(a-u)}{a^4} & u < a \\ 0 & u \geq a \end{cases} \qquad \text{(linear)} \tag{9a}$$

$$\rho(u) = (a^3/8\pi)e^{-au} \qquad \text{(exponential)} \tag{9b}$$

with associated dipole−dipole interaction tensors:

$$T_{ij} = (4v^3 - 3v^4)\delta_{ij}/r^3 - 3v^4\,(r_ir_j/r^5) \qquad r < s$$

$$s = a\,(\alpha_q\alpha_p)^{1/6} \qquad v = r/s \qquad \text{(linear)} \tag{9c}$$

$$T_{ij} = \delta_{ij}/r^3 - 3\,r_ir_j/r^5 \qquad r > s$$

and

$$T_{ij} = \delta_{ij}/r^3[1 - (a^2r^2/2 + ar + 1)e^{-ar}] - 3r_ir_j/r^5[1 - (a^3r^3/6 + a^2r^2/2 + ar + 1)e^{-ar}] \qquad \text{(exponential)} \tag{9d}$$

By taking the isotropic atomic polarizabilities (for H, C, N, and O, regardless of their bonding situation) and the screening length a as parameters, these models were fitted to 16 experimental molecular polarizabilities. For six unrelated molecules, which were not in the learning set, the polarizabilities were reproduced with experimental accuracy. In the meantime we have found ample evidence about the transferability of the parameters, and the model also works well if, e.g., Hartree−Fock atomic polarizabilities are used for atoms beyond the set parametrized by Thole.[10]

TABLE 8: CPHF Percentage of Experimental Value

group	performance (%)	group	performance (%)
alcohols	77.83 ± 5.56	diatomics	58.80 ± 10.17
alkanes	81.62 ± 4.15	halogens	70.56 ± 7.68
alkenes	78.38 ± 5.08	sulfurs	69.83 ± 1.54
carbonyls	76.62 ± 5.51	various	69.97 ± 10.02
cyanides	81.68 ± 6.03	average	73.47 ± 10.08
		Halogens	
F	67.77 ± 2.19	Br	78.82 ± 8.96
Cl	64.18 ± 1.65	I	75.92 ± 5.67

Recent studies have shown the success of this in QM/MM calculations (direct reaction field approach) involving discrete classical solvent molecules to account for the near surroundings, and a dielectric medium for the long-distance interactions. It has been used for a number of chemical problems: dimers of benzene derivatives,[4] S_N1 dissociation of *tert*-butyl chloride,[11] solvatochromic behavior of acetone,[12,13] solvent-induced dipole moments of the excited states of twisted ethylene,[14] solvent effects on (hyper)polarizabilities,[15] etc.

Thole presented a set of polarizabilities for just H, C, N, and O, but since then more and different (i.e., halogen-containing) molecules have been used in calculations and simulations. Hence, there was a need for more model atomic polarizabilities. To enable reasonable comparison of QM/MM calculations with completely QM or completely classical calculations, we needed also parameters for modeling ab initio polarizabilities.

Empirical Polarizabilities

Experimental data are usually obtained from the Lorentz−Lorentz equation:

$$\frac{n^2-1}{n^2+2}\frac{M}{\rho}=\frac{4}{3}\pi N_a\bar{\alpha} \qquad (10)$$

which relates the refractive index n (usually at 5893 Å, sodium D line) with the mean molecular polarizability $\bar{\alpha}$. This equation is accurate in the gas state, but an extrapolation of the refractive index to zero frequency (or infinite wavelength) should be made:

$$n(\lambda)=n^*_{\infty}+\frac{a}{\lambda^2}+\frac{b}{\lambda^4}+... \qquad (11)$$

As a result of this extrapolation, the polarizabilities decrease within 2−4%,[16] which gives an estimate of the uncertainty in the experimental data. The values obtained at 5893 Å are used throughout this article, but with this uncertainty kept in mind.

Another source of molecular polarizabilities are the Kerr constants, which give an estimate of both the mean polarizability and its anisotropy. However, in using these constants, an assumption has to be made of the geometry of the molecule, thereby reducing the validity to small (symmetric) molecules. And even then, this assumption introduces an additional uncertainty.

Ab Initio Polarizabilities

When there are no experimental data available for a given molecule, ab initio values can be used, e.g., in QM/MM calculations. They also provide a more reasonable comparison between the QM/MM and fully quantum mechanical results when effective atomic polarizabilities are used to represent the molecular polarizability in the basis set used in the fully QM calculations.

TABLE 9

molecule	exp	STO	N31	DZV	DZP	TZV	TZP
Second-Order Perturbation Polarizabilities (au)							
acetone	43.12	14.89	26.32	26.97	29.38	29.47	30.90
acetonitrile	30.23	11.42	18.51	19.53	21.15	20.53	21.98
carbon monoxide	13.16	5.16	8.01	8.29	9.05	9.00	9.44
cyclohexane	74.23	24.81	49.22	50.97	54.44	54.17	56.83
cyclopentane	61.75	20.28	40.03	42.27	45.26	44.91	47.23
dimethyl ether	35.36	10.33	18.67	18.69	21.13	21.10	22.70
p-dioxane	58.04	20.22	35.64	35.27	39.38	39.47	41.83
ethane	30.23	8.49	16.99	17.79	19.47	18.91	20.36
ethanol	34.28	10.16	18.85	19.30	21.70	21.22	22.90
formaldehyde	16.53	6.24	9.95	9.78	10.96	10.94	11.56
hydrogen	5.33	0.90	1.65	1.65	1.99	1.69	2.03
methanol	22.40	6.07	10.59	10.70	12.50	11.99	13.21
nitrogen	11.88	5.37	8.27	4.13	5.76	8.85	9.35
oxygen	10.80	3.87	5.60	5.58	6.87	5.95	6.60
propane	43.05	12.63	25.40	26.68	28.89	28.52	30.35
tert-butyl cyanide	64.72	24.11	43.59	45.60	48.99	48.99	51.58
acetamide	38.26	13.71	22.12	22.96	25.50	25.45	26.95
ethylene oxide	29.89	10.45	17.71	18.02	19.92	19.72	20.91
formamide	27.53	9.61	14.39	14.69	16.55	16.67	17.71
methane	17.68	4.39	8.76	9.11	10.20	9.69	10.68
propanol	47.04	14.30	27.09	27.95	30.93	30.50	32.60
water	10.06	1.88	2.85	3.63	4.17	3.60	4.38
Unsöld Polarizabilities (au)							
acetone	43.12	28.50	35.20	37.25	32.11	37.24	33.21
acetonitrile	30.23	21.14	25.80	27.65	23.58	27.53	24.20
carbon monoxide	13.16	11.16	12.22	13.20	10.26	13.22	10.68
cyclohexane	74.23	45.69	61.05	64.22	57.70	64.13	59.13
cyclopentane	61.75	37.71	50.13	53.31	48.06	52.94	49.16
dimethyl ether	35.36	21.12	27.11	27.99	23.69	28.26	24.78
p-dioxane	58.04	39.40	48.23	49.75	42.80	49.75	44.35
ethane	30.23	16.73	23.14	24.72	21.54	24.68	22.10
ethanol	34.28	20.90	26.96	28.49	24.16	28.54	25.05
formaldehyde	16.53	12.73	14.85	15.57	12.55	15.67	13.14
hydrogen	5.33	1.06	1.66	1.66	2.46	1.70	2.50
methanol	22.40	13.35	16.72	17.53	14.38	17.70	14.99
nitrogen	11.88	11.35	12.51	9.24	7.28	12.89	10.74
oxygen	10.80	9.52	9.36	9.90	7.89	9.93	7.96
propane	43.05	24.35	33.35	35.42	31.34	35.43	32.24
tert-butyl cyanide	64.72	44.69	56.62	59.77	52.87	59.58	54.44
acetamide	38.26	27.03	31.25	33.62	28.39	33.87	29.62
ethylene oxide	29.89	20.58	24.57	25.84	22.04	25.72	22.72
formamide	27.53	19.87	21.38	22.95	18.86	23.54	19.99
methane	17.68	9.11	13.15	14.09	11.88	14.11	12.22
propanol	47.04	28.55	37.18	39.22	33.88	39.25	35.03
water	10.06	5.97	5.94	6.54	5.11	7.37	5.55
Coupled Perturbative HF Polarizabilities (au)							
acetone	43.12	16.80	30.81	31.99	33.34	34.74	35.54
acetonitrile	30.23	11.35	20.80	22.23	23.25	23.62	24.59
carbon monoxide	13.16	4.41	8.59	9.29	9.73	10.22	10.43
cyclohexane	74.23	31.24	57.43	59.22	60.75	62.97	64.19
cyclopentane	61.75	25.63	47.36	49.50	50.82	52.45	53.60
dimethyl ether	35.36	13.21	24.02	24.05	25.50	26.70	27.62
p-dioxane	58.04	25.07	42.78	42.92	45.16	47.07	48.29
ethane	30.23	10.82	21.58	22.33	23.23	23.83	24.65
ethanol	34.28	13.07	23.87	24.44	25.82	26.64	27.58
formaldehyde	16.53	6.14	11.54	11.73	12.52	13.02	13.52
hydrogen	5.33	1.00	2.12	2.11	2.53	2.19	2.62
methanol	22.40	7.73	13.90	14.10	15.25	15.63	16.38
nitrogen	11.88	3.94	7.76	−	8.82	8.76	9.26
oxygen	10.80	2.86	5.69	5.99	7.03	6.53	7.14
propane	43.05	16.09	31.38	32.56	33.72	34.86	35.87
tert-butyl cyanide	64.72	27.60	50.30	52.78	54.59	56.52	57.98
acetamide	38.26	15.12	25.90	27.46	28.95	30.26	31.08
ethylene oxide	29.89	11.81	21.24	21.93	22.85	23.56	24.17
formamide	27.53	9.98	16.66	17.64	18.82	19.96	20.59
methane	17.68	5.54	11.59	11.96	12.60	12.79	13.40
propanol	47.04	18.43	33.61	34.56	36.24	37.46	38.60
water	10.06	2.39	3.90	4.20	5.09	5.01	5.56

TABLE 10: Experimental and Fitted Polarizabilities (au) for the Halogen-Containing Molecules. Polarizabilities for H and C, and *a* as in Table 2

model	exp	linear	dev, %	expon	dev, %
fluoromethane	17.61	18.09	2.69	18.30	3.89
chloromethane	30.57	30.90	1.09	31.27	2.28
bromomethane	37.45	37.45	0.00	37.45	0.00
difluoromethane	18.42	18.43	0.06	18.56	0.74
dichloromethane	45.89	44.48	3.07	44.30	3.46
fluoroform	18.90	18.80	0.53	18.80	0.52
chloroform	57.36	58.29	1.62	57.19	0.30
carbon tetrafluoride	19.23	19.15	0.45	19.03	1.07
carbon tetrachloride	70.86	71.82	1.36	69.37	2.09
chlorine	31.11	30.66	1.44	32.09	3.13
average dev, %			1.23		1.75

There are several ab initio methods to calculate polarizabilities, which are all based on a Taylor expansion of the total energy around the electric field strength **F**

$$U = U^{(0)} - \mu_i^{(0)} F_i - \frac{1}{2!}\alpha_{ij} F_i F_j - \frac{1}{3!}\beta_{ijk} F_i F_j F_k - \frac{1}{4!}\gamma_{ijkl} F_i F_j F_k F_l \ldots \qquad (12)$$

where the subscripts correspond to the tensor components which are summed over the Cartesian axes *x*, *y*, and *z*. $U^{(0)}$ is energy of the unperturbed system, F_i is the component of **F** in the *i*th direction, $\mu^{(0)}$ is the permanent dipole moment, $\boldsymbol{\alpha}$ is the static dipole polarizability, and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are the first and second dipole hyperpolarizability tensors, respectively.

For any wave function, the finite field (FF) method is applicable.[17] For this paper we computed polarizabilities using the second-order perturbation (SOP) method, extended with the nonempirical Unsöld approximation[18,19] (USOP), and the coupled perturbative Hartree−Fock method (CPHF).[20] Both the CPHF and FF methods take into account the actual response of a molecule to an applied electric field and give, therefore, in principle the same results, but each has its own characteristics. The CPHF method is analytical and therefore faster, but it is limited to closed-shell RHF wave functions. The FF method is numerical, and therefore slower, but can be used with any wave function.

The SOP expression for the polarizability is:

$$\alpha_{ij} = -2\sum_{k\neq 0}^{\infty} \frac{\langle\Psi_0|\mu^i|\Psi_k\rangle\langle\Psi_k|\mu^j|\Psi_0\rangle}{E_0 - E_k} \qquad (13a)$$

in which Ψ_0 and Ψ_k represent the system's ground and *k*th excited states, respectively, E_0 and E_k are the corresponding state energies, and μ^i is the *x*, *y*, or *z* component of the dipole moment operator.

In the Unsöld approximation(s),[21] the excitation energies in the denominator are replaced by a single parameter, Δ, usually an experimental excitation or the ionizaton energy. Thus, after applying the closure relation, the polarizability is reduced to the difference between two ground-state expectation values:

$$\alpha_{ij} = -\frac{2}{\Delta}\{\sum_{k=0}^{\infty}\langle\Psi_0|\mu^i|\Psi_k\rangle\langle\Psi_k|\mu^j|\Psi_0\rangle - \langle\Psi_0|\mu^i|\Psi_0\rangle\langle\Psi_0|\mu^j|\Psi_0\rangle\}$$

$$= -\frac{2}{\Delta}\{\langle\mu^i\mu^j\rangle - \langle\mu^i\rangle\langle\mu^j\rangle\} \qquad (13b)$$

In the Møller−Plesset approximation,[22] which, like CPHF, can only be applied for wave functions in which orbital energies are meaningful, this is simplified by replacing the *state* functions by molecular *orbitals*, and the energies by the corresponding *orbital* energies. For a closed-shell determinant wave function this leads to:

$$\alpha_{ij} = -4\sum_{p}^{\text{occ}}\sum_{q}^{\text{virt}} \frac{\mu_{pq}^i \mu_{pq}^j}{\epsilon_p - \epsilon_q} \qquad (13c)$$

A good choice for Δ_{ij} for a finite basis was shown to be:[18,19]

$$\frac{1}{\Delta_{ij}} = \frac{\displaystyle\sum_{k\neq 0}^{n} \frac{\langle 0|\mu^i|k\rangle\langle k|\mu^j|0\rangle}{E_0 - E_k}}{\displaystyle\sum_{k\neq 0}^{n}\langle 0|\mu^i|k\rangle\langle k|\mu^j|0\rangle} \qquad (14)$$

leading to the *nonempirical* Unsöld approximation, which is nearly basis set independent. Applying the closure relation, which is also an approximation for a finite basis, turns eq 12 into:

$$\alpha_{ij}^u = -2\sum_{k\neq 0}^{n}\frac{\langle 0|\mu^i|k\rangle\langle k|\mu^j|0\rangle}{E_0 - E_k}\left[\frac{\langle 0|\mu^i\mu^j|0\rangle - \langle 0|\mu^i|0\rangle\langle 0|\mu^j|0\rangle}{\displaystyle\sum_{k\neq 0}^{n}\langle 0|\mu^i|k\rangle\langle k|\mu^j|0\rangle}\right] = \alpha_{ij}C_{ij} \qquad (15a)$$

where C_{ij} can be interpreted as a factor correcting for the incompleteness of the basis set.

For a Hartree−Fock wave function we have:

$$\langle 0|\mu^i\mu^j|0\rangle = 2\sum_{p}^{\text{occ}} Q_{pp}^{ij} + \sum_{p,q}^{\text{occ}}\mu_{pp}^i\mu_{qq}^i - \sum_{p,q}^{\text{occ}}\mu_{pq}^i\mu_{pq}^i \qquad (15b)$$

as follows from the standard rules for the evaluation of matrix elements for one- and two-electron operators between determinantal wave functions. Here the Q_{pq}^{ij} is the *i*,*j* component of the *p*,*q* matrix element of the nontraceless Cartesian second moment operator.

After correcting α_{ij} of eq 12a with C_{ij}, $\boldsymbol{\alpha}^u$ is nearly basis set independent. However, $\boldsymbol{\alpha}^u$ is not a proper tensor because C lacks the correct transformation properties, but for average polarizabilities $\boldsymbol{\alpha}^u$ is sufficient. These procedures have been implemented in the HONDO8.1[20] program.

Computational Details

All calculations were carried out on a HP cluster and the Cray J932 supercomputer in Groningen. We used the HONDRF[5] program, which is essentially the HONDO8.1 package, extended with (among many other things) the finite field procedures. We used MOPAC93[23] for some geometry optimizations and the new program AOPT[24] for fitting the polarizabilities.

We started by repeating Thole's work, i.e., the same learning set and the same set of test cases, to check the performance of the new procedure and find the effect of possibly different molecular geometries. These geometries were optimized with HONDO using a closed-shell RHF wave function in Dunning's double-ζ valence (DZV)[25] basis set.

Thole originally minimized the relative difference between experimental and calculated polarizability components, but

TABLE 11: Polarizabilities Fitted to Experimental and ab Initio Values (Molecules in *Italics* Were Not in the Learning Set)

model		exp	to exp		CPHF	to CPHF	
			linear	expon		linear	expon
Alcohols							
2-propanol	$CH_3CHOHCH_3$	47.04	48.39	47.26	37.72	37.83	36.75
ethanol	C_2H_5OH	34.28	35.92	35.23	26.86	26.86	26.49
methanol	CH_3OH	22.40	22.77	22.64	15.42	17.13	15.92
cyclohexanol	$C_6H_{11}OH$	*78.01*	*80.29*	*77.97*	*65.48*	*64.29*	*63.40*
dev, %			3.06	1.09	22.17	3.30	2.59
Alkanes							
cyclohexane	C_6H_{12}	74.23	75.08	73.02	63.42	60.95	60.87
cyclopentane	C_5H_{10}	61.75	62.68	61.26	52.87	50.90	50.87
cyclopropane	C_3H_6	38.06	37.09	38.02	32.06	30.36	32.06
ethane	C_2H_6	30.23	30.13	29.95	23.60	23.60	23.96
hexane	C_6H_{14}	79.63	83.02	79.84	66.22	67.82	67.22
methane	CH_4	17.86	17.52	17.68	12.95	13.33	13.54
propane	C_3H_8	43.05	42.89	42.19	34.58	34.14	34.32
dodecane	$C_{12}H_{26}$	*153.86*	*168.12*	*157.76*	*131.44*	*140.00*	*136.16*
neopentane	$C(CH_3)_4$	*68.83*	*67.17*	*65.49*	*54.93*	*54.27*	*54.11*
dev, %			2.64	1.57	18.38	3.03	2.36
Alkenes							
benzene	C_6H_6	70.18	61.06	61.87	56.99	50.42	54.90
chlorobenzene	C_6H_5Cl	82.67	75.47	75.60	66.44	58.75	63.28
ethylene	C_2H_4	28.48	25.28	25.93	21.49	20.03	21.78
nitrobenzene	$C_6H_5NO_2$	87.19	76.91	76.53	72.83	61.47	63.98
acetylene	C_2H_2	*23.55*	*20.34*	*21.90*	*15.83*	*16.81*	*19.87*
m-dichlorobenzene	$C_6H_4Cl_2$	*96.03*	*91.24*	*89.65*	*77.46*	*69.34*	*71.54*
o-dichlorobenzene	$C_6H_4Cl_2$	*95.62*	*90.23*	*89.03*	*76.75*	*68.41*	*71.26*
dev, %			9.86	8.87	21.61	10.43	8.89
Carbonyls							
N-methylformamide	$HCONHCH_3$	39.88	39.47	39.32	30.67	31.22	30.25
acetaldehyde	$HCOCH_3$	30.97	30.44	30.70	23.61	23.36	24.04
acetamide	CH_3CONH_2	38.26	39.37	38.57	31.95	29.95	29.45
acetone	CH_3COCH_3	43.12	43.49	43.14	34.84	34.42	34.62
formaldehyde	$HCOH$	16.53	17.76	18.25	13.04	13.26	13.39
formamide	$HCONH_2$	27.53	26.19	26.53	19.28	19.10	19.28
N,N-dimethylformamide	$HCON(CH_3)_2$	*52.70*	*52.53*	*51.77*	*41.76*	*43.10*	*40.81*
N-methylacetamide	$CH_3CONHCH_3$	*52.77*	*52.74*	*51.95*	*41.90*	*42.04*	*40.94*
carbonyl chloride	$COCl_2$	*45.75*	*44.27*	*44.30*	*29.55*	*28.95*	*28.58*
dev, %			2.49	2.63	23.38	2.06	2.46
Cyanides							
ethyl cyanide	C_2H_5CN	42.11	41.97	42.24	34.62	34.06	36.32
methyl cyanide	CH_3CN	30.23	29.17	29.84	24.02	23.54	25.64
methyl dicyanide	$CH_2(CN)_2$	39.07	41.07	42.31	34.87	34.11	38.48
tert-butyl cyanide	$(CH_3)_3CCN$	64.72	67.27	66.33	56.79	55.60	56.57
chloromethyl cyanide	CH_2ClCN	*41.16*	*42.66*	*42.87*	*31.94*	*31.50*	*33.20*
isopropyl cyanide	$(CH_3)_2CHCN$	*54.32*	*54.93*	*54.54*	*46.21*	*45.24*	*46.65*
trichloromethyl cyanide	CCl_3CN	*70.32*	*69.56*	*68.26*	*49.53*	*47.03*	*48.27*
dev, %			2.68	2.84	18.32	2.35	4.26
Diatomics							
carbon monoxide	CO	13.16	13.16	13.16	9.73	10.26	10.33
chlorine	Cl_2	31.11	30.71	31.20	18.02	14.76	18.02
hydrogen	H_2	5.33	5.21	4.90	2.58	2.58	2.58
hydrogen bromide	HBr	24.36	24.51	25.96	14.08	15.00	17.29
hydrogen chloride	HCl	17.75	17.70	18.45	8.35	9.32	10.45
hydrogen iodide	HI	36.78	36.78	39.65	21.30	23.19	25.14
nitrogen	N_2	11.88	11.88	11.88	8.73	8.73	8.73
nitric oxide	NO	11.74	11.34	11.36	5.37	7.43	7.00
oxygen	O_2	10.80	10.80	10.80	7.23	5.87	5.10
dev, %			0.87	3.32	41.20	11.97	14.66
Halogens							
bromomethane	CH_3Br	37.86	37.71	37.86	27.55	27.33	27.55
chloromethane	CH_3Cl	30.70	30.83	30.70	20.09	21.55	20.98
dibromomethane	CH_2Br_2	58.58	58.61	57.99	42.31	42.31	41.80
diiodomethane	CH_2I_2	87.05	84.41	82.54	72.06	54.48	56.68
fluoromethane	CH_3F	17.68	18.15	18.15	12.69	13.12	13.29
iodomethane	CH_3I	51.22	50.18	51.22	35.29	35.29	35.29
tetrachloromethane	CCl_4	70.92	70.92	68.46	45.29	44.64	43.35
tetrafluoromethane	CF_4	19.71	19.37	19.35	13.04	12.94	12.62
trichloromethane	$CHCl_3$	57.56	57.91	56.31	37.38	37.38	35.96
trifluoromethane	CHF_3	18.96	18.96	18.96	12.77	12.93	12.83
dichloromethane	CH_2Cl_2	*46.02*	*44.26*	*43.55*	*28.12*	*29.93*	*28.44*
difluoromethane	CH_2F_2	*18.42*	*18.54*	*18.56*	*12.53*	*13.00*	*13.06*
tribromomethane	$CHBr_3$	*79.90*	*79.90*	*75.99*	*73.11*	*50.76*	*55.12*
trichlorofluoromethane	$CFCl_3$	*55.61*	*57.64*	*56.45*	*36.45*	*36.77*	*35.68*
triiodomethane	CHI_3	*121.74*	*121.73*	*111.74*	*92.65*	*73.12*	*77.35*
dev, %			1.27	2.47	29.43	6.80	6.14

TABLE 11 (continued)

			to exp			to CPHF	
model		exp	linear	expon	CPHF	linear	expon
			Sulfurs				
carbon disulfide	CS_2	58.98	44.82	41.46	39.92	36.92	38.05
sulfur dioxide	SO_2	26.25	26.25	26.25	18.54	20.63	19.29
sulfur hexafluoride	SF_6	30.17	37.77	33.09	21.48	21.48	21.48
Percentage			16.40	13.13	30.16	6.26	2.91
			Various				
ammonia	NH_3	14.98	13.38	13.17	8.81	8.67	8.44
carbon dioxide	CO_2	17.75	17.45	18.74	12.39	12.39	13.39
dimethyl ether	CH_3OCH_3	35.36	35.95	35.36	26.59	30.08	26.66
ethylene oxide	CH_2OCH_2	29.90	29.90	30.52	23.88	24.62	23.88
p-dioxane	$(CH_2)_4O_2$	58.04	60.50	59.29	46.70	49.75	46.44
water	H_2O	10.06	10.08	10.06	5.22	5.34	5.14
nitrous oxide	N_2O	*20.24*	*17.19*	*17.81*	*14.94*	*11.88*	*11.98*
Percentage			4.79	4.84	30.02	6.73	4.92

because of the additional uncertainty in the experimental components, we minimized the relative difference in the mean polarizabilities instead:

$$\Delta = \sum_{i}^{N} \frac{|\bar{\alpha}_{\text{group},i} - \bar{\alpha}_{\text{exp},i}|}{\bar{\alpha}_{\text{exp},i}} \qquad (16)$$

Using the two shape functions, the associated atomic polarizabilities were optimized at a given *a* value, and subsequently the *a* values were optimized.

Next, we computed the polarizabilities for these molecules using six of the basis sets available in HONDO: minimal basis set (STO-3G), double-ζ valence 4-31G (N31), Dunning's double-ζ valence (DZV), double-ζ valence plus polarization functions on all atoms (DZP), triple-ζ valence (TZV), and triple-ζ valence plus polarization (TZP),[25] with the SOP, USOP, and CPHF methods. The resulting polarizabilities were used as input to AOPT, giving the "Thole parameters" for each basis.

Finally, we extended the learning set to 52 molecules, in particular with halogenides, and defined a new test set, comprising 18 molecules. The experimental geometries of the diatomics were taken from Applequist et al.[26] For the remaining molecules we took the experimental geometries as listed in the manual to MOPAC93[27] or were (if not available there) optimized with MOPAC(SCF/PM3).

In addition, for all 70 molecules ab initio polarizabilities were calculated. In general the DZP basis was used here. For Br- and I-containing molecules we used effective core potentials (ECPs),[28] which are claimed to account for relativistic effects, together with uncontracted (i.e., n s- and n p-) valence, diffuse valence, and polarization functions.[29] The Thole parameters were obtained for the CPHF polarizabilities, or FF for molecules where an open-shell wave function is appropriate.

Results and Discussion

In Table 1 and Table 2 the results of fitting the experimental average polarizabilities of the original learning and control sets are collected. The small average deviation in the mean polarizabilities for the learning set is well within the experimental uncertainty, while the new Thole parameters have become somewhat different in comparison with the original values (Table 1, column Thole 2). The smaller average deviation in the mean polarizability is apparent and could be caused by either the different geometries or by the fit procedure. We therefore applied Thole's original parameters to the present geometries (Table 1, column Thole 1), giving virtually the old results. The differences are therefore almost completely caused by the new fit procedure.

In Table 3 we report the (CPHF) calculated polarizabilities of the control set in the various basis sets, and in Table 4 the corresponding parameters. In Table 9 more complete information is collected. Although the deviations here are somewhat larger than for the fit to the experimental polarizabilities, they are still within experimental error bounds, and the parameters are therefore quite useful for generating classical polarizabilities matching the basis set selected.

Some additional information about the basis set dependence, in combination with the method, of the polarizabilities (expressed as the average fraction of the experimental values) is presented in Table 5. Obviously, one could not expect too much from the STO-3G basis set. But even there the nonempirical Unsöld approximation, for the mean polarizability, leads to values surprisingly close to those from larger basis sets and from more sophisticated methods. The general conclusion from Table 5 is that it is advisable to use the USOP method whenever possible, e.g., with N31 or DZV, to obtain passable average polarizabilities at minimal computational cost.

As a first extension to our learning set, we chose a number of halogen-containing molecules in order to find atomic polarizabilities for these atoms. The *a* values and H, C, N, and O atomic polarizabilities were kept fixed. Again the resulting mean polarizabilities were close to experiment (see Table 10) with polarizabilities (Table 6), which make physical sense in relation to the polarizabilities of the other atoms.

Finally, the learning set was extended to 52 molecules, divided into nine groups (with a not-too-strict nomenclature): alcohols, alkanes, "alkenes", "carbonyls", cyanides, diatomics, halogenides, "sulfurs", and "various". Except for the "sulfurs", at least one molecule from each group was put into the new control set. Thole parameters calculated from both experimental and ab initio CPHF/DZP (except for Br and I; see Computational Details) mean polarizabilities are given in Table 8, the pertinent data in Table 11. Since these calculations provide the components of the polarizabilities proper for the actual molecular geometry, another optimization was done, now fitting these components. The associated parameters are also contained in Table 8, the full list of data in Table 12.

In the contents of Table 7, we note that the effective atomic polarizabilities are very close to those in Tables 2 and 6, as are the screening parameters. This supports our experience, i.e., that the "old" parameters have "universal" meaning in the sense that the model predicts molecular polarizabilities almost regardless of the chemical composition.

Finally, in Table 8 we give the percentage of the experimental mean polarizabilities obtained from the CPHF calculations in the DZP basis. Generally one gets at best about 80% from this

TABLE 12: Polarizabilities Fitted to Calculated Mean and Components (au)

	CPHF				linear				expon			
	α_{mean}	α_{xx}	α_{yy}	α_{zz}	α_{mean}	α_{xx}	α_{yy}	α_{zz}	α_{mean}	α_{xx}	α_{yy}	α_{zz}
Alcohols												
2-propanol	37.72	37.72	37.63	37.80	36.79	39.96	35.28	35.12	36.51	38.61	34.41	36.52
ethanol	26.86	29.38	25.28	25.91	26.77	33.20	22.60	24.50	26.41	31.54	23.81	23.87
methanol	15.42	15.93	16.21	14.13	17.06	20.90	15.74	14.53	15.97	15.60	17.01	15.30
cyclohexanol	*65.48*	*68.24*	*32.11*	*66.10*	*61.04*	*65.97*	*53.38*	*63.78*	*62.42*	*67.04*	*55.42*	*64.79*
dev, %					5.05				3.28			
Alkanes												
cyclohexane	63.42	65.10	64.35	60.79	57.93	62.17	60.28	51.34	59.97	63.00	61.65	55.25
cyclopentane	52.87	54.36	54.45	49.80	48.67	52.87	53.51	39.64	50.16	52.81	53.24	44.43
cyclopropane	32.06	33.32	33.32	29.54	30.00	32.81	32.81	24.37	33.11	35.12	35.12	29.10
ethane	23.60	24.68	23.06	23.06	23.07	28.72	20.24	20.24	24.25	27.32	22.72	22.72
hexane	66.22	72.86	65.76	60.04	65.24	85.76	61.78	48.17	67.10	82.07	64.88	54.34
methane	12.95	12.95	12.95	12.95	13.78	13.78	13.77	13.78	13.91	13.91	13.91	13.91
propane	34.58	33.23	33.48	37.03	33.11	27.62	29.99	41.70	34.47	30.82	32.33	40.26
dodecane	*131.44*	*149.89*	*130.53*	*113.91*	*134.32*	*182.42*	*129.79*	*90.74*	*134.95*	*170.27*	*131.57*	*103.00*
neopentane	*54.93*	*54.93*	*54.93*	*54.93*	*52.62*	*52.62*	*52.62*	*52.62*	*53.85*	*53.84*	*53.85*	*53.85*
dev, %					4.87				3.37			
Alkenes												
benzene	56.99	71.95	71.95	27.06	46.92	57.90	57.90	24.97	55.58	70.84	70.83	25.08
chlorobenzene	66.44	79.24	88.25	31.85	56.33	65.40	73.79	29.82	61.31	75.33	79.32	29.27
ethylene	21.49	34.87	19.61	9.98	18.84	24.86	20.11	11.54	22.81	34.10	22.88	11.44
nitrobenzene	72.83	90.53	94.88	33.10	58.58	69.99	75.99	29.76	62.39	77.84	80.29	29.04
acetylene	*15.83*	*30.02*	*8.73*	*8.73*	*15.72*	*25.07*	*11.04*	*11.04*	*23.10*	*48.83*	*10.23*	*10.23*
m-dichlorobenzene	*77.46*	*93.01*	*103.05*	*36.33*	*66.64*	*78.21*	*86.73*	*34.98*	*66.60*	*81.44*	*84.65*	*33.71*
o-dichlorobenzene	*76.75*	*92.43*	*101.52*	*36.28*	*65.80*	*77.75*	*84.89*	*34.74*	*66.55*	*81.50*	*84.57*	*33.56*
dev, %					13.39				14.84			
Carbonyls												
N-methylformamide	30.67	38.88	31.81	21.32	30.83	39.92	31.46	21.11	28.76	34.19	30.78	21.32
acetaldehyde	23.61	24.08	28.51	18.24	22.86	22.71	28.39	17.47	24.07	24.56	29.52	18.14
acetamide	31.95	32.10	35.56	28.19	29.08	32.17	30.35	24.73	27.84	29.53	30.72	23.26
acetone	34.84	36.66	39.80	29.08	34.23	41.99	33.13	27.57	34.48	40.74	35.17	27.52
formaldehyde	13.04	18.64	13.20	7.27	13.24	16.15	15.17	8.41	13.21	16.99	14.57	8.05
formamide	19.28	20.74	26.25	10.85	18.15	21.05	22.72	10.68	18.03	20.35	23.26	10.49
N,N-dimethylformamide	*41.76*	*43.99*	*49.35*	*31.93*	*43.35*	*47.41*	*50.68*	*31.97*	*39.03*	*41.50*	*44.44*	*31.15*
N-methylacetamide	*41.90*	*43.85*	*49.42*	*32.45*	*41.79*	*46.06*	*48.66*	*30.64*	*39.39*	*44.52*	*42.90*	*30.73*
carbonyl chloride	*29.55*	*33.27*	*38.88*	*16.50*	*30.29*	*30.68*	*41.92*	*18.26*	*23.61*	*27.04*	*26.85*	*16.95*
dev, %					3.16				6.94			
Cyanides												
ethyl cyanide	34.62	29.80	44.41	29.65	32.70	26.91	41.05	30.15	35.94	28.32	49.71	29.80
methyl cyanide	24.02	34.31	18.87	18.87	22.84	31.28	18.62	18.62	25.22	39.80	17.93	17.93
methyl dicyanide	34.87	27.04	29.88	47.70	32.48	25.63	27.68	44.11	37.73	25.80	29.27	58.13
tert-butyl cyanide	56.79	65.67	52.58	52.12	54.11	58.06	52.23	52.03	55.39	65.61	50.54	50.01
chloromethyl cyanide	*31.94*	*36.91*	*25.74*	*33.18*	*30.75*	*35.96*	*25.34*	*30.94*	*30.31*	*27.65*	*26.74*	*36.53*
isopropyl cyanide	*46.21*	*56.36*	*42.83*	*39.45*	*43.89*	*51.42*	*46.08*	*34.17*	*45.77*	*58.86*	*42.85*	*35.60*
trichloromethyl cyanide	*49.53*	*48.04*	*49.06*	*51.49*	*46.85*	*46.09*	*46.77*	*47.67*	*40.59*	*36.52*	*38.49*	*46.77*
dev, %					5.17				6.23			
Diatomics												
carbon monoxide	9.73	12.78	8.20	8.20	9.45	11.94	8.20	8.20	9.31	13.54	4.38	4.38
chlorine	18.02	33.02	10.53	10.53	15.90	21.21	13.24	13.24	13.58	16.73	12.01	12.01
hydrogen	2.58	6.39	0.67	0.67	1.36	2.12	0.98	0.98	2.88	4.55	2.04	2.04
hydrogen bromide	14.08	17.41	12.42	12.42	13.96	16.28	12.80	12.80	14.18	17.30	12.62	12.62
hydrogen chloride	8.35	12.97	6.04	6.04	8.61	10.42	7.70	7.70	8.34	10.59	7.21	7.21
hydrogen iodide	21.30	26.21	18.85	18.85	21.58	24.08	20.33	20.33	22.04	25.70	20.21	20.21
nitrogen	8.73	7.93	6.49	11.78	8.08	10.51	6.86	6.86	6.17	9.06	4.72	4.72
nitric oxide	5.37	5.93	4.43	5.76	7.10	9.79	5.75	5.75	5.44	7.56	4.38	4.38
oxygen	7.23	13.91	3.54	4.23	6.02	9.36	4.34	4.34	4.73	6.23	3.97	3.97
dev, %					13.73				12.23			
Halogens												
bromomethane	27.55	38.30	22.14	22.21	27.15	38.30	21.57	21.57	24.47	28.14	22.63	22.63
chloromethane	20.09	27.57	16.35	16.35	21.67	29.80	17.61	17.61	18.81	20.70	17.86	17.86
dibromomethane	42.31	32.26	34.11	60.56	41.88	31.04	33.16	61.45	35.36	31.44	32.12	42.52
diiodomethane	72.06	44.41	48.95	122.81	57.11	44.75	50.27	76.32	51.15	44.37	47.59	61.48
fluoromethane	12.69	12.36	12.86	12.86	13.48	12.35	14.05	14.05	13.66	12.31	14.34	14.34
iodomethane	35.29	28.01	28.03	49.82	34.92	28.11	28.11	48.54	32.19	29.39	29.39	37.80
tetrachloromethane	45.29	45.29	45.29	45.29	45.29	45.29	45.29	45.29	34.18	34.18	34.18	34.18
tetrafluoromethane	13.04	13.04	13.04	13.04	13.23	13.23	13.23	13.23	13.33	13.33	13.33	13.33
trichloromethane	37.38	26.09	43.02	43.02	37.27	27.66	42.07	42.07	28.94	26.09	30.37	30.37
trifluoromethane	12.77	12.47	12.91	12.91	13.23	13.87	12.91	12.91	13.41	14.40	12.91	12.91
dichloromethane	*28.12*	*34.47*	*29.92*	*19.97*	*29.73*	*36.08*	*31.41*	*21.70*	*23.82*	*25.41*	*24.28*	*21.78*
difluoromethane	*12.53*	*12.49*	*12.63*	*12.47*	*13.33*	*12.73*	*13.01*	*14.24*	*13.53*	*12.75*	*13.07*	*14.77*
tribromomethane	*73.11*	*42.07*	*100.10*	*77.14*	*53.96*	*43.41*	*59.23*	*59.23*	*46.17*	*41.68*	*48.42*	*48.42*
trichlorofluoromethane	*36.45*	*28.34*	*40.50*	*40.50*	*37.30*	*28.18*	*41.86*	*41.86*	*28.94*	*26.01*	*30.40*	*30.40*
triiodomethane	*92.65*	*56.48*	*111.11*	*110.36*	*76.90*	*57.83*	*86.44*	*86.44*	*69.78*	*58.47*	*75.43*	*75.43*
dev, %					6.76				15.95			

TABLE 12 (continued)

	CPHF				linear				expon			
	α_{mean}	α_{xx}	α_{yy}	α_{zz}	α_{mean}	α_{xx}	α_{yy}	α_{zz}	α_{mean}	α_{xx}	α_{yy}	α_{zz}
					Sulfurs							
carbon disulfide	39.92	82.24	18.76	18.76	30.25	47.49	21.63	21.63	33.87	58.38	21.61	21.61
sulfur dioxide	18.54	25.28	18.05	12.28	17.84	23.08	18.14	12.28	16.56	20.50	16.85	12.32
sulfur hexafluoride	21.48	21.48	21.48	21.48	18.14	18.14	18.14	18.14	21.42	21.42	21.42	21.42
dev, %					14.52				8.70			
					Various							
ammonia	8.81	10.04	10.04	6.34	8.13	9.20	8.24	8.24	7.55	8.53	8.53	5.58
carbon dioxide	12.39	21.68	7.74	7.74	11.86	17.86	8.85	8.85	13.18	23.47	8.04	8.04
dimethyl ether	26.59	29.85	25.53	24.39	30.71	41.89	25.84	24.39	26.43	30.00	25.41	23.89
ethylene oxide	23.88	30.14	18.32	23.18	24.84	31.19	20.84	22.49	23.85	30.29	17.21	24.05
p-dioxane	46.70	42.63	53.79	43.67	48.90	46.66	59.31	40.72	45.56	44.53	49.72	42.44
water	5.22	6.53	6.09	3.03	4.65	5.68	5.23	3.03	5.32	6.39	5.86	3.70
nitrous oxide	*14.94*	*27.05*	*8.18*	*9.59*	*11.50*	*18.01*	*8.24*	*8.24*	*9.35*	*15.45*	*6.30*	*6.30*
dev, %					10.02				9.03			

basis set on the SCF level. For diatomics and planar systems usually results are worse, because on needs here additional diffuse and/or eccentric basis functions to improve the off-axis (diatomics) or the off-plane (planar molecules) components of the polarizabilities.

Thole's model suffers the same shortcomings since it is based on *isotropic* atoms, and any anisotropy in the molecular polarizability comes from the interaction tensors, which are, of course, only defined by the atom positions. In principle, this situation can be improved by adding off-axis or off-plane (input) polarizabilities as was tried for benzene, but it involves more parameters to determine (additional polarizabilities and their positions), and it leads in practice to rather bad interaction energies.[4]

Conclusion

We reviewed Thole's modified polarizability model and extended it to a wider set of effective atomic polarizabilities. The introduction of "new" atoms and the redefinition of the learning and control sets affected the original parameters only slightly, while the "new" polarizabilities (halogens and sulfur) fit nicely with what one would expect from their atomic numbers. The model can be used to predict molecular polarizabilities from the geometry. It provides a simple, consistent, and independent way for defining atomic polarizabilities in "polarizable force fields" and in QM/MM applications.

Acknowledgment. This paper is dedicated to the memory of Theo Thole who died on 4 July 1996 after a fatal accident.

References and Notes

(1) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.

(2) Gao, J. *J. Phys. Chem.* **1995**, *99*, 16460.

(3) Thompson, M. A. *J. Phys. Chem.* **1996**, *100*, 14492.

(4) van Duijnen, P. Th.; de Vries, A. H. *Int. J. Quant. Chem.* **1996**, *60*, 1111.

(5) de Vries, A. H.; van Duijnen, P. Th.; Juffer, A. H.; Rullmann, J. A. C.; Dijkman, J. P.; Merenga, H.; Thole, B. T. *J. Comput. Chem.* **1995**, *16*, 37.

(6) Rullmann, J. A. C.; Bellido, M. N.; van Duijnen, P. Th. *J. Mol. Biol.* **1989**, *206*, 101.

(7) van Duijnen, P. Th.; de Vries, A. H. *Int. J. Quant. Chem. Quant. Chem. Symp.* **1995**, *29*, 523.

(8) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, J. S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(9) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.

(10) Bellido, M. N. *Chem. Phys. Lett.* **1985**, *122*, 562.

(11) Swart, M.; de Vries, A. H.; van Duijnen, P. Th. Manuscript i, preparation.

(12) de Vries, A. H.; van Duijnen, P. Th. *Int. J. Quant. Chem.* **1996**, *57*, 1067.

(13) Grozema, F.; van Duijnen, P. Th. Submitted for publication.

(14) Zijlstra, R. W. J.; van Duijnen, P. Th.; de Vries, A. H. *Chem. Phys.* **1996**, *204*, 439.

(15) van Duijnen, P. Th.; Swart, M.; Grozema, F. QM/MM calculation of (hyper-) polarizabilities with the DRF approach. *ACS Symposium Series*; American Chemical Society: Washington, DC, 1998 (in press).

(16) Böttcher, C. J. F.; Bordewijk, P. *Theory of electric polarization*; Elsevier: Amsterdam, 1978; Vol. II.

(17) Sim, F.; Chin, S.; Dupuis, M.; Rice, J. E. *J. Phys. Chem.* **1993**, *97*, 158.

(18) Mulder, F.; van Hemert, M. C.; Wormer, P. E. S.; van der Avoird, A. *Theor. Chim. Acta* **1977**, *46*, 39.

(19) van Hemert, M. C.; Blom, C. E. *Mol. Phys.* **1981**, *43*, 229.

(20) Dupuis, M.; Farazdel, A.; Karma, S. P.; Maluendes, S. A. HONDO: a general atomic and molecular electronic structure system. In *MOTECC-90*; Clementi, E., Ed.; ESCOM: Leiden, 1990; p 277.

(21) Unsöld, A. *Z. Phys.* **1927**, *43*, 563.

(22) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *146*, 618.

(23) Stewart, J. J. P. *MOPAC93*; Fujitsu Ltd.: Tokyo, Japan, 1993.

(24) Swart, M. *AOPT*; University of Groningen: Groningen, The Netherlands, 1996.

(25) Dunning, T. H., Jr.; Hay, P. J. Gaussian basis sets for molecular, calculations. In *Methods of Electronic Structure Theory*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977; p 1.

(26) Applequist, J.; Carl, J. R.; Fung, J. K. *J. Am. Chem. Soc.* **1972**, *94*, 2952.

(27) Stewart, J. J. P. *MOPAC 93.00 Manual*; Fujitsu Ltd.: Tokyo, 1993.

(28) Bergner, A.; Dolg, M.; Küchle, W.; Stoll, H.; Preuss, H. *Mol. Phys.* **1993**, *80*, 1431.

(29) GLukhovstev, M. N.; Pross, A.; McGrath, M. P.; Radom, L. *J. Chem. Phys.* **1995**, *103*, 1878.